L. Kantorovitch, G. Akilov

# ANALYSE FONCTIONNELLE

## Tome 2

## Equations fonctionnelles

Éditions Mir Moscou

Л. В. КАНТОРОВИЧ, Г. П. АКИЛОВ

# ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ

ИЗДАТЕЛЬСТВО « НАУКА »
МОСКВА

L. KANTOROVITCH, G. AKILOV

# ANALYSE FONCTIONNELLE

Tome 2

Equations fonctionnelles

ÉDITIONS MIR • MOSCOU

Traduit du russe par
*Djilali Embarek*

*На французском языке*

# RAPPEL:

## SOMMAIRE DU TOME 1

# TABLE DES MATIÈRES

CHAPITRE XII

# ÉQUATION ADJOINTE

## § 1. Théorèmes de l'opérateur inverse *)

Dans ce paragraphe, nous complétons ce que nous avons dit sur l'opérateur inverse dans le tome 1 (cf. § 4, chap. V).

1.1. Rappel des définitions du tome 1. Soit $U$ un opérateur linéaire continu d'un espace normé X dans un espace normé Y. Si existe un opérateur $V$ de Y dans X, tel que

$$VU = I_{\mathbf{X}} \quad (I_{\mathbf{X}} x = x, \quad x \in \mathbf{X}), \tag{1}$$

$$UV = I_{\mathbf{Y}} \quad (I_{\mathbf{Y}} y = y, \quad y \in \mathbf{Y}), \tag{2}$$

on dit que $V$ est inverse de $U$ ($V = U^{-1}$).

Dire qu'il existe un opérateur inverse (même non continu) $U^{-1}$ revient à dire que $U$ réalise une bijection de X sur Y. Si $U^{-1}$ est continu, l'application indiquée est un isomorphisme.

L'opérateur $V$ est appelé opérateur à gauche (resp. à droite) et noté $V = U_g^{-1}$ (resp. $V = U_d^{-1}$ **)) si est réalisée la relation (1) (resp. (2)). Dans V.4.4, on a montré qu'une condition nécessaire et suffisante pour qu'existe un opérateur inverse à gauche continu est que

$$\| U(x) \| \geqslant m \| x \| \quad (x \in \mathbf{X}), \tag{3}$$

où $m > 0$ ne dépend pas de $x$. Si, en outre, l'opérateur $U$ applique X *sur* Y, l'opérateur inverse à gauche sera également inverse à droite, c'est-à-dire l'opérateur inverse continu $U^{-1}$ existe.

1.2. Prouvons un théorème.

THÉORÈME 1. *Si un opérateur linéaire continu $U$ d'un B-espace* X *dans un espace normé* Y *admet un inverse à gauche continu, l'ensemble* $\mathbf{Y}' = U(\mathbf{X})$ *est un B-espace.*

DÉMONSTRATION. Il faut simplement établir la complétude de l'espace Y'. Soit $\{y_n\}$ une suite de Cauchy d'éléments de Y'. Posons $x_n = U^{-1}(y_n)$ ($y_n = U(x_n)$; $n = 1, 2, \ldots$). D'après ce qu'on a dit

---

*) Ce sujet est également traité dans Neumann [1] et Schauder [2].

**) On omettra parfois les indices « $g$ » et « $d$ ».

dans 1.1

$$\| y_n - y_k \| = \| U(x_n) - U(x_k) \| \geqslant m \| x_n - x_k \|,$$

où $m$ est une constante positive. Donc

$$\lim_{k,\, n\to\infty} \| x_n - x_k \| = 0.$$

Par suite, dans X existe l'élément $x_0 = \lim\limits_{n\to\infty} x_n$. Comme $\lim y_n = \lim U(x_n) = U(x_0)$, en désignant $y_0 = U(x_0)$ on obtient $y_0 \in Y'$ et $y_n \to y_0$. Ce qui prouve la complétude de Y'.

COROLLAIRE. *Sous les hypothèses du théorème*, Y' *est un sous-espace fermé de* Y.

**1.3.** Soient donnés deux espaces normés X et Y et un opérateur linéaire continu $U$ de X dans Y. L'ensemble $X_0 = U^{-1}(0)$ est manifestement un sous-espace fermé de X. Considérons l'espace quotient $\overline{X} = X/X_0$ (cf. IV.1.8, tome 1). Soit $\overline{x} \in \overline{X}$; prenons un élément quelconque $x \in \overline{x}$ et posons

$$\overline{U}(\overline{x}) = U(x). \tag{4}$$

La définition de l'élément $\overline{U}(\overline{x})$ ne dépend pas du choix de l'élément $x \in \overline{x}$, car si $x', x'' \in \overline{x}$, alors $x' - x'' \in X_0$ et $U(x') = U(x'')$. Donc, la formule (4) définit un opérateur $\overline{U}$ de $\overline{X}$ dans Y. Cet opérateur est homogène et additif. Il est continu, car en passant à la borne inférieure dans le second membre de l'inégalité

$$\| \overline{U}(\overline{x}) \| = \| U(x) \| \leqslant \| U \| \, \| x \| \qquad (x \in \overline{x}),$$

on peut écrire

$$\| \overline{U}(\overline{x}) \| \leqslant \| U \| \, \| \overline{x} \| \qquad (\overline{x} \in \overline{X}).$$

De la définition de l'opérateur $\overline{U}$ il suit que $U = \overline{U}\varphi$, où $\varphi$ est un homomorphisme naturel de X sur $X/X_0$ (cf. IV. 1.8, tome 1) et $\| U \| = \| \overline{U} \|$.

Contrairement à $U$ l'opérateur $\overline{U}$ réalise une application injective (de $\overline{X}$ dans Y). En effet, si $\overline{U}(\overline{x}) = 0$, pour tout $x \in \overline{x}$ on a $U(x) = 0$, c'est-à-dire $x \in U^{-1}(0) = X_0$ et par suite $\overline{x}$ est confondu avec $X_0$, l'élément nul de $\overline{X}$.

Si $U$ applique X sur Y, $\overline{U}$ appliquera également $\overline{X}$ sur Y. Si, de plus, existe l'opérateur inverse continu $\overline{U}^{-1}$, les espaces X et Y sont dits *homomorphes* et l'opérateur $U$, *homomorphisme* de X sur Y.

Les deux conditions suivantes expriment le fait que $U$ est un homomorphisme de X sur Y:

1) $U(X) = Y$;

2) il existe $m > 0$ tel que pour tout $y \in Y$ on peut exhiber $x \in X$ tel que

$$y = U(x), \quad \| y \| \geqslant m \| x \|.$$

En effet, si $U$ est un homomorphisme, la première condition est manifeste. D'autre part, en prenant $x \in \bar{x} = \bar{U}^{-1}(y)$ à partir de (4) de IV.1.8 (tome 1), on obtient

$$\| x \| \leqslant 2 \| \bar{x} \| \leqslant 2 \| \bar{U}^{-1} \| \| y \|$$

et l'on peut prendre $m = \dfrac{1}{2 \| \bar{U}^{-1} \|}$.

Si, au contraire, les deux conditions sont remplies, de la première on déduit que $\bar{U}(\bar{X}) = Y$. Soit $\bar{x} \in \bar{X}$; à partir de l'élément $y = \bar{U}(\bar{x})$ trouvons l'élément $x \in X$ vérifiant la deuxième condition. Comme

$$\bar{U}(\varphi(x)) = U(x) = y = \bar{U}(\bar{x}),$$

et que l'opérateur réalise une bijection $\bar{U}$, on a $\varphi(x) = \bar{x}$ et par suite

$$\| \bar{U}(\bar{x}) \| = \| y \| \geqslant m \| x \| \geqslant m \| \bar{x} \|.$$

Comme indiqué dans 1.1, ceci et la relation $\bar{U}(\bar{X}) = Y$ assurent l'existence de l'opérateur inverse continu $U^{-1}$.

**1.4.** La réciproque du théorème 1 est capitale en théorie des équations fonctionnelles. C'est une partie de la proposition suivante.

LEMME 1. *Soit $U$ un opérateur linéaire continu d'un $B$-espace* X *dans un espace normé* Y. *Si l'image $U(B)$ de la boule unité $B$ (de centre* 0) *de l'espace* X *est dense dans la boule $S_r$ de rayon $r$ (de centre* 0 *aussi) de l'espace* Y, *alors $U$ est un homomorphisme de* X *sur* Y. *En particulier, si l'application réalisée par l'opérateur $U$ est bijective, l'opérateur $U$ admet un opérateur inverse continu $U^{-1}$.*

DÉMONSTRATION. Vérifions que les deux conditions de 1.3 sont remplies. De toute évidence, on peut admettre que les boules $B$ et $S_r$ sont fermées; montrons que

$$U(B) \supset S_{r/2}. \tag{5}$$

Prenons une suite $\{\varepsilon_n\}$ de nombres strictement positifs, telle que $\sum_{k=1}^{\infty} \varepsilon_k \leqslant 1$ et considérons $y \in S_r$. Comme $\overline{U(B)} \supset S_r$, il existe $y_1 \in U(B)$ tel que

$$\| y - y_1 \| \leqslant \varepsilon_1 r.$$

Soient $x_1$ un élément de $B$, tel que $y_1 = U(x_1)$, $B_h$ la boule fermée de X de rayon $h$ (de centre 0). Par hypothèse, $\overline{U(B_h)} \supset S_{hr}$. Comme

$y - y_1 \in S_{\varepsilon_1 r}$, il existe donc un élément $x_2 \in B_{\varepsilon_1}$, tel que

$$\| y - (y_1 + y_2) \| \leqslant \varepsilon_2 r \quad (y_2 = U(x_2)).$$

En poursuivant ce raisonnement, on trouve deux suites $\{y_n\} \subset \mathrm{Y}$ et $\{x_n\} \subset \mathrm{X}$ telles que

$$y_n = U(x_n), \quad x_n \in B_{\varepsilon_{n-1}}, \quad \| y - \sum_{k=1}^{n} y_k \| \leqslant \varepsilon_n r \tag{6}$$

$$(n \in \mathbb{N}, \quad \varepsilon_0 = 1).$$

Comme $\| x_n \| \leqslant \varepsilon_{n-1}$ et que l'espace X est complet, la série $\sum_{k=1}^{\infty} x_k$ est convergente. Si $x$ est la somme de cette série, on a

$$\| x \| \leqslant \sum_{k=1}^{\infty} \| x_k \| \leqslant \sum_{k=1}^{\infty} \varepsilon_{k-1} \leqslant 2,$$

c'est-à-dire $x \in B_2$. D'autre part

$$U(x) = \sum_{k=1}^{\infty} U(x_k) = \sum_{k=1}^{\infty} y_k.$$

Mais de (6) il est clair que $\sum_{k=1}^{\infty} y_k = y$. Donc, $y = U(x)$. On a ainsi prouvé que $U(B_2) \supset S_r$, ce qui équivaut à (5).

Comme (5) entraîne que $U(B_n) \supset S_{nr/2}$, on a

$$U(\mathrm{X}) = \bigcup_{n=1}^{\infty} U(B_n) \supset \bigcup_{n=1}^{\infty} S_{nr/2} = \mathrm{Y},$$

et la première condition est remplie.

Si $y \neq 0$ est un élément quelconque de Y, alors

$$y' = \frac{r}{2\| y \|} y \in S_{r/2}$$

et, d'après (5), on peut exhiber un élément $x' \in B$, tel que $y' = U(x')$. En posant $x = \frac{2\| y \|}{r} x'$, on obtient

$$U(x) = y, \quad \| x \| = \frac{2}{r} \| y \| \| x' \| \leqslant \frac{2}{r} \| y \|.$$

Donc, la deuxième condition est aussi satisfaite. C.q.f.d.

Corollaire. *Si l'on se place dans les conditions du lemme, l'espace* Y *est complet.*

En effet, $U$ étant un homomorphisme, l'opérateur $\overline{U}$ de l'espace quotient $\overline{\mathrm{X}} = \mathrm{X}/\mathrm{X}_0$ $(\mathrm{X}_0 = U^{-1}(0))$ sur Y admet un inverse continu. En appliquant le théorème 1, on obtient ce qu'on voulait, puisque $\mathrm{Y} = \overline{U}(\overline{\mathrm{X}})$.

L'hypothèse du lemme est de vérification délicate. Celle du théorème suivant est plus commode.

THÉORÈME 2 (Banach). *Si l'ensemble $U$ (X) est de deuxième catégorie dans l'espace* Y, *l'hypothèse du lemme est satisfaite, donc l'opérateur $U$ est un homomorphisme de* X *sur* Y.

DÉMONSTRATION. Conservons les notations du lemme et prouvons que, si l'hypothèse du lemme n'est pas remplie, l'ensemble $U(B)$ n'est nulle part dense. En effet, en supposant le contraire on trouve une boule $S(y_0, r)$ (de centre $y_0$ et de rayon $r$) dans l'espace Y, telle que

$$\overline{U(B)} \supset S(y_0, r). \tag{7}$$

L'ensemble $U(B)$ est symétrique, c'est-à-dire contient tout élément $y$ avec son opposé $-y$. L'adhérence $\overline{U(B)}$ est visiblement symétrique, elle aussi. Donc, en vertu de (7), on peut écrire

$$\overline{U(B)} \supset S(-y_0, r).$$

Soit un élément $y \in S_r$. L'élément $y_0 + y \in S(y_0, r)$ et l'élément $-y_0 + y \in S(-y_0, r)$, donc ces deux éléments appartiennent à $\overline{U(B)}$. Or, l'ensemble $U(B)$, donc $\overline{U(B)}$, est convexe, par suite il contient deux quelconques de ses éléments avec leur demi-somme. Notamment

$$y = \frac{(y_0+y)+(-y_0+y)}{2} \in \overline{U(B)}.$$

Donc, $\overline{U(B)} \supset S_r$.

Si la condition du lemme 1 n'est pas remplie, l'ensemble $U(B)$ n'est nulle part dense. Ce sera également le cas pour tout ensemble $U(B_n)$ $(n \in \mathbb{N})$. Or

$$U(\mathrm{X}) = \bigcup_{n=1}^{\infty} U(B_n),$$

d'où il suit que l'ensemble $U$ (X) est de première catégorie. Cette contradiction prouve le théorème.

Indiquons un corollaire du théorème prouvé, qui est la réciproque du théorème 1.

COROLLAIRE. *Si un opérateur linéaire continu $U$ réalise une application bijective d'un $B$-espace* X *sur un sous-espace fermé d'un $B$-espace* Y, *l'opérateur inverse $U^{-1}$ est continu.*

En effet, tout sous-espace fermé d'un $B$-espace est un $B$-espace, donc un ensemble de deuxième catégorie dans soi (cf. I.4.7, tome 1).

1.5. Indiquons quelques applications immédiates du théorème 2.

Supposons qu'un espace vectoriel X soit muni d'une norme de deux manières différentes. Soient $\|x\|_1$ et $\|x\|_2$ les normes respecti-

ves d'un élément $x \in X$, $X_1$ et $X_2$ les espaces normés correspondants. Même si $X_1$ et $X_2$ doivent être considérés comme des espaces différents, il peut ne pas exister de distinctions qualitatives entre eux. C'est notamment le cas lorsque toute suite $\{x_n\}$ convergente dans un espace est convergente dans l'autre vers le même élément. On dit alors que les normes des espaces $X_1$ et $X_2$ sont équivalentes ; ce qui veut dire que les espaces $X_1$ et $X_2$ sont isomorphes (cf. IV.1.3, tome 1).

THÉORÈME 3. *Soient* $X_1$ *et* $X_2$ *deux B-espaces normés tels que* $X_1 \subset \subset X_2$ *). *Si la convergence* $x_n \to x$ *dans l'espace* $X_1$ *entraîne la convergence* $x_n \to x$ *dans l'espace* $X_2$, *alors ou bien* $X_1 = X_2$ *et les normes de* $X_1$ *et* $X_2$ *sont équivalentes, ou bien* $X_1$ *est un ensemble de première catégorie dans* $X_2$.

DÉMONSTRATION. Désignons par $U$ un opérateur d'immersion de $X_1$ dans $X_2$, c'est-à-dire un opérateur associant à un élément $x \in X_1$ cet élément $x$ lui-même, mais considéré comme élément de $X_2$. D'après le théorème, l'opérateur $U$ est un opérateur linéaire continu. Si l'ensemble $X_1 = U(X_1)$ est de deuxième catégorie dans l'espace $X_2$, alors $X_1 = X_2$ d'après le théorème 2 et $U$ admet un opérateur inverse continu. Donc, si $x_n \to x$ dans $X_2$, alors $x_n = U^{-1}(x_n) \to \to U^{-1}(x) = x$ dans $X_1$, c'est-à-dire les normes de $X_1$ et $X_2$ sont équivalentes.

Si dans le théorème $X_1 = C^{(1)}(D)$ et $X_2 = C(D)$, l'ensemble de toutes les fonctions continûment différentiables est de première catégorie dans l'espace C des fonctions continues. On s'assure de façon analogue que l'ensemble des fonctions bornées presque partout mesurables est de première catégorie dans l'espace $L^1$, etc.

1.6. Une application $T$ (pas nécessairement linéaire) d'un ensemble $\Omega$ d'un espace normé X dans un espace normé Y est par définition *fermée* si

$$x_n \in \Omega \quad (n = 1, 2, \ldots), \quad x_n \to x_0, \quad T(x_n) \to y_0,$$

entraîne que $x_0 \in \Omega$ et $T(x_0) = y_0$.

Un opérateur linéaire continu défini sur un ensemble fermé est visiblement fermé. La réciproque est vraie. De façon plus précise, on a le

THÉORÈME 4. *Si* $T$ *est un opérateur linéaire fermé d'un sous-espace vectoriel fermé* $\Omega$ *d'un B-espace* X *dans un B-espace* Y, *c'est que* $T$ *est continu.*

DÉMONSTRATION. On peut admettre que $\Omega = X$ (puisque $\Omega$ est un $B$-espace). Munissons X d'une nouvelle norme en posant

$$\| x \|_1 = \| x \| + \| T(x) \| \quad (x \in X). \tag{8}$$

---

*) On admet que l'immersion $X_1 \subset X_2$ conserve les opérations algébriques, c'est-à-dire $X_1$ peut être considéré comme un sous-espace vectoriel de $X_2$.

Il est immédiat de vérifier que cette norme satisfait aux axiomes d'un espace normé. Vérifions que l'espace X est complet pour cette nouvelle norme. Soit

$$\lim_{k,\, n\to\infty} \| x_n - x_k \|_1 = 0.$$

Ceci exprime que

$$\lim_{k,\, n\to\infty} \| x_n - x_k \| = 0 \quad \text{et} \quad \lim_{k,\, n\to\infty} \| T(x_n) - T(x_k) \| = 0.$$

Les espaces X (pour la norme donnée) et Y étant complets, il résulte qu'existent les limites

$$\lim_{n\to\infty} x_n = x \quad \text{et} \quad \lim_{n\to\infty} T(x_n) = y_0.$$

L'opérateur $T$ étant fermé, on a $y_0 = T(x_0)$. Or

$$\lim_{n\to\infty} \| x_n - x_0 \|_1 = \lim_{n\to\infty} \| x_n - x_0 \| + \lim_{n\to\infty} \| T(x_n) - T(x_0) \| = 0$$

ce qui prouve que l'espace X est complet pour la nouvelle norme.

Comme

$$\| x \| \leqslant \| x \|_1,$$

de $\| x_n \|_1 \to 0$ il suit que $\| x_n \| \to 0$. En appliquant le théorème précédent on obtient

$$\| x \|_1 \leqslant M \| x \|.$$

*A fortiori*

$$\| Tx \| \leqslant M \| x \|,$$

ce qui exprime la continuité de l'opérateur $T$.

Remarque. La classe des opérateurs linéaires fermés définis sur l'espace tout entier (ou sur un sous-espace vectoriel fermé) est confondue avec la classe des opérateurs linéaires continus. Cependant, si l'on considère les opérateurs linéaires fermés sur un sous-espace vectoriel (non fermé), ils forment une classe bien plus large que celle des opérateurs continus. Ainsi, dans l'espace $L^2(a, b)$ on vérifie immédiatement que l'opérateur $T$:

$$y = T(x), \quad y(t) = \frac{dx(t)}{dt},$$

défini sur l'ensemble $\Omega$ de toutes les fonctions absolument continues, dont les dérivées premières appartiennent à $L^2(a, b)$, est fermé mais pas continu.

On a affaire aux opérateurs fermés mais non continus essentiellement dans le cas où X = Y est un espace hilbertien. L'étude du cas général est rendue malaisée par la structure compliquée de l'espace de Banach.

## § 2. Relation entre une équation et son adjointe

Dans ce paragraphe on considère l'équation

$$U(x) = y \tag{1}$$

et l'équation

$$U^*(g) = f, \tag{2}$$

dite *adjointe* de (1). On suppose toujours que $U$ est un opérateur linéaire continu d'un espace X dans un espace Y. Sans le mentionner expressément dans la suite, nous admettrons que ces espaces sont complets, bien que certains théorèmes de ce paragraphe soient valables sans l'hypothèse de complétude.

Les théorèmes de ce paragraphe ont été prouvés par Hellinger et Toeplitz [1], [2] pour l'espace $\mathbf{L}^2(a, b)$, par Riesz [3] pour les espaces $\mathbf{l}^p$ et $\mathbf{L}^p(a, b)$ et par Banach et Zaanen [I] pour le cas général.

2.1. Les ensembles d'annulation auront un grand rôle à jouer dans la suite de l'exposé. Soit $\Gamma$ un ensemble de fonctionnelles linéaires dans l'espace X. Désignons par $\mathbf{N}(\Gamma)$ l'ensemble de tous les $x \in \mathbf{X}$ tels que $f(x) = 0$ pour tout $f \in \Gamma$. Si $E \subset \mathbf{X}$ nous appelons $\mathbf{N}^*(E)$ l'ensemble de toutes les fonctionnelles $f \in \mathbf{X}^*$ s'annulant sur chaque élément de $E$. Les ensembles $\mathbf{N}(\Gamma)$ et $\mathbf{N}^*(E)$ sont dits *ensembles d'annulation* de $\Gamma$ et $E$ respectivement.

Si nous considérons le couple en dualité $\langle \mathbf{X}, \mathbf{X}^* \rangle$, il est évident que $\mathbf{N}(\Gamma)$ est l'annulateur $\Gamma^\perp$ de l'ensemble $\Gamma$, et $\mathbf{N}^*(E)$ l'annulateur $E^\perp$ de l'ensemble $E$, introduits dans la situation générale au III.3.2 (voir tome 1). En tenant compte des propriétés des annulateurs et des polaires, développées au III.3.2, et du fait que la fermeture en norme et la fermeture faible sont confondues dans X, on obtient

1) $\mathbf{N}(\Gamma)$ est un sous-espace vectoriel fermé de X;
2) $\mathbf{N}(\mathbf{N}^*(E))$ est l'enveloppe linéaire fermée de $E$;
3) si $\mathbf{X}_0$ est un sous-espace fermé de X, alors

$$\mathbf{X}_0 = \mathbf{N}(\mathbf{N}^*(\mathbf{X}_0)).$$

De façon analogue

1) $\mathbf{N}^*(E)$ est un sous-espace vectoriel (*)-faiblement fermé de $\mathbf{X}^*$;
2) $\mathbf{N}^*(\mathbf{N}(\Gamma))$ est l'enveloppe linéaire (*)-faiblement fermée de $\Gamma$;
3) si Z est un sous-espace (*)-faiblement fermé de $\mathbf{X}^*$, alors

$$\mathbf{Z} = \mathbf{N}^*(\mathbf{N}(\mathbf{Z})).$$

Les sous-espaces de $\mathbf{X}^*$, fermés pour la norme, n'étant pas (*)-faiblement fermés si X n'est pas réflexif, les sous-espaces fermés ne sont pas tous de la forme $\mathbf{N}^*(E)$ $(E \subset \mathbf{X})$.

Soit $\pi$: $\mathbf{X} \to \mathbf{X}^{**}$ une injection canonique. Si $E$ est un ensemble de X, on désigne $\pi(E)$ par $E_0^{**}$, c'est-à-dire $E_0^{**}$ est l'ensemble de toutes les fonctionnelles de $\mathbf{X}^*$ de la forme $F_x$ $(x \in \mathbf{X})$, où comme

toujours

$$F_x(f) = \overline{f(x)} \quad (f \in \mathbf{X}^*).$$

De façon analogue, on se servira de la notation $G_y$ ($y \in \mathbf{Y}$) pour l'espace $\mathbf{Y}$.

On a manifestement

$$\mathbf{N}^*(E) = \mathbf{N}(E_0^{**}).$$

Pour l'ensemble $\Gamma \subset \mathbf{X}^*$, on peut considérer deux ensembles d'annulation : $\mathbf{N}(\Gamma) \subset \mathbf{X}$ et $\mathbf{N}^*(\Gamma) \subset \mathbf{X}^{**}$. Il est immédiat de voir que

$$[\mathbf{N}(\Gamma)]_0^{**} = \mathbf{N}^*(\Gamma) \cap \mathbf{X}_0^{**},$$

où $\mathbf{X}_0^{**}$ est l'ensemble $\pi(\mathbf{X})$ dans les notations adoptées.

Introduisons enfin l'ensemble $\mathbf{N}(U)$, ou *ensemble d'annulation de l'opérateur* $U$, c'est-à-dire l'ensemble de tous les $x \in \mathbf{X}$, dont l'image par $U$ est zéro. Autrement dit, $\mathbf{N}(U) = U^{-1}(0)$.

2.2. Le théorème suivant précise les éléments qui constituent l'image $U(\mathbf{X})$ de l'espace $\mathbf{X}$.

THEOREME 1. *Soient* $\mathbf{Y}' = U(\mathbf{X})$, $\mathbf{Y}_1$ *l'adhérence de* $\mathbf{Y}'$. *Alors*

$$\mathbf{Y}_1 = \mathbf{N}(\mathbf{N}(U^*)),$$

*c'est-à-dire* $\mathbf{Y}_1$ *est l'ensemble des zéros communs des fonctionnelles* $g \in \mathbf{Y}^*$ *sur lesquelles s'annule l'opérateur* $U^*$.

DEMONSTRATION. Prouvons tout d'abord la relation

$$\mathbf{N}^*(\mathbf{Y}') = \mathbf{N}(U^*). \tag{3}$$

Soit $g \in \mathbf{N}^*(\mathbf{Y}')$. Comme $U(x) \in \mathbf{Y}'$ pour tout $x \in \mathbf{X}$, on a

$$U^*(g)(x) = g(U(x)) = 0 \quad (x \in \mathbf{X}), \tag{4}$$

c'est-à-dire $U^*(g) = 0$ et $g \in \mathbf{N}(U^*)$. Cette même égalité (4) montre que si $U^*(g) = 0$, alors $g \in \mathbf{N}^*(\mathbf{Y}')$. Donc, la relation (3) est établie. En passant aux ensembles d'annulation dans les deux membres de cette relation, on obtient

$$\mathbf{N}(\mathbf{N}^*(\mathbf{Y}')) = \mathbf{N}(\mathbf{N}(U^*)).$$

$\mathbf{Y}'$ étant un espace vectoriel, son adhérence, comme indiqué dans 2.1, est confondue avec $\mathbf{N}(\mathbf{N}^*(\mathbf{Y}'))$. C.q.f.d.

On a le théorème dual.

THEOREME 1*. *Si* $\tilde{\mathbf{X}}^*$ *est l'adhérence* (*)-*faible de l'ensemble* $U^*(\mathbf{Y}^*)$, *alors*

$$\tilde{\mathbf{X}}^* = \mathbf{N}^*(\mathbf{N}(U)).$$

DEMONSTRATION. Dans la relation (3), remplaçons l'opérateur $U$ par $U^*$ :

$$\mathbf{N}^*(U^*(\mathbf{Y}^*)) = \mathbf{N}(U^{**})$$

et considérons l'intersection des deux membres avec l'ensemble $X_0^{**}$. Comme

$$U^{**}(F_x)(g) = F_x(U^*(g)) = g(U(x)) = G_{U(x)}(g)$$
$$(g \in Y^*, \quad x \in X),$$

c'est-à-dire

$$U^{**}(F_x) = G_{U(x)} \quad (x \in X),$$

l'intersection $N(U^{**}) \cap X_0^{**}$ est composée de toutes les fonctionnelles $F_x \in X_0^{**}$ telles que $x \in N(U)$. En d'autres termes

$$N(U^{**}) \cap X_0^{**} = [N(U(Y^*))]_0^{**}.$$

Par ailleurs

$$N^*(U^*(Y^*)) \cap X_0^{**} = [N(U^*(Y^*))]_0^{**}.$$

En combinant les relations obtenues on trouve

$$N(U^*(Y^*)) = N(U),$$

et par suite

$$N^*(N(U^*(Y^*))) = N^*(N(U)).$$

Or $\tilde{X}^* = N^*(N(U^*(Y^*)))$, ce qui prouve le théorème.

2.3. Les deux théorèmes suivants établissent un lien entre le fait que l'une des équations (1) ou (2) admet une solution pour tout second membre, et le fait que l'autre admet une solution unique.

THÉORÈME 2. *Pour que l'équation* (2) *admette une solution pour tout* $f \in X^*$, *il est nécessaire et suffisant que l'opérateur U admette un inverse à gauche continu* $U^{-1}$.

DÉMONSTRATION. *La condition est nécessaire.* Supposons que l'équation (2) admet une solution pour tout $f \in X^*$. En d'autres termes, soit $U^*(Y^*) = X^*$. Considérons un élément quelconque $x \neq 0$ dans l'espace X. D'après le théorème V.7.2 (tome 1) il existe une fonctionnelle $f \in X^*$ telle que

$$\| f \| = 1, \quad f(x) = \| x \|.$$

Le théorème 1.2 nous dit que $U^*$ est un homomorphisme, donc il existe $g \in Y^*$ telle que

$$U^*(g) = f, \quad \| g \| \leqslant m \| f \|,$$

où $m$ ne dépend pas de $f$ et par suite de $x$. D'après ce qui précède

$$\| x \| = f(x) = g(U(x)) \leqslant \| g \| \, \| U(x) \| \leqslant m \, \| U(x) \|.$$

Donc

$$\| U(x) \| \geqslant \frac{1}{m} \| x \| \quad (x \in X),$$

d'où résulte l'existence de l'opérateur inverse à gauche continu $U^{-1}$.

*La condition est suffisante.* Soit $f_0 \in \mathbf{X}^*$. Supposons qu'existe l'opérateur continu $U_g^{-1} = U^{-1}$ et posons

$$g'(y) = f_0(U^{-1}(y)) \quad (y \in U(\mathbf{X})).$$

La fonctionnelle $g'$ définie sur l'ensemble $\mathbf{Y}' = U(\mathbf{X})$ est linéaire et continue. En la prolongeant à $\mathbf{Y}$ tout entier, on obtient la fonctionnelle $g_0 \in \mathbf{Y}^*$. Comme $U(x) \in \mathbf{Y}'$ pour tout $x \in \mathbf{X}$, on obtient

$$U^*(g_0)(x) = g_0(U(x)) = g'(U(x)) = f_0(U^{-1}U(x)) = f_0(x).$$

Donc $U^*g = f_0$, c'est-à-dire l'équation (2) admet une solution pour tout second membre.

Théorème 2*. *Pour que l'équation* (1) *admette une solution pour tout* $y \in \mathbf{Y}$, *il est nécessaire et suffisant que l'opérateur* $U^*$ *possède un inverse à gauche continu* ${U^*}^{-1}$.

Démonstration. *Condition nécessaire.* On l'établit comme celle du théorème 2. Soit $y \in \mathbf{Y}$ un élément quelconque de norme $\| y \| \leqslant 1$. D'après le théorème 1.2, il existe dans $\mathbf{X}$ un élément $x$ tel que

$$U(x) = y, \quad \| x \| \leqslant m \| y \| \leqslant m.$$

Si $g \in \mathbf{Y}^*$ et $f \in U^*(g)$, alors

$$| g(y) | = | g(U(x)) | = | f(x) | \leqslant \| f \| \| x \| \leqslant m \| f \|.$$

En passant à la borne supérieure sur $y$ au premier membre ($\| y \| \leqslant 1$) et en tenant compte du fait que $m$ ne dépend pas de $y$, on obtient

$$\| g \| = \sup_{\| y \| \leqslant 1} | g(y) | \leqslant m \| f \| = m \| U^*(g) \|,$$

et nous arrivons à la condition assurant l'existence de l'opérateur inverse à gauche continu ${U^*}^{-1}$.

*Condition suffisante.* Montrons que l'opérateur $U$ vérifie l'hypothèse du lemme 1 du paragraphe précédent et, par suite, en vertu du lemme, $U(\mathbf{X}) = \mathbf{Y}$.

Soit $B$ la boule unité de l'espace $\mathbf{X}$ (centrée en 0). L'ensemble $U(B)$ est absolument convexe, donc son adhérence dans l'espace $\mathbf{Y}$ est $U(B)^{\circ\circ}$ (cf. III.3.2, tome 1). Le polaire $U(B)^\circ$ est l'ensemble de toutes les fonctionnelles $g \in \mathbf{Y}^*$ telles que $| g(y) | \leqslant 1$ pour tous les $y \in U(B)$. Ou encore

$$| g(U(x)) | \leqslant 1 \quad \text{pour tous les} \quad x \in B.$$

En posant $f = U^*g$, on peut mettre la dernière relation sous la forme

$$| f(x) | \leqslant 1 \quad (x \in B).$$

Donc, lorsque $g$ parcourt l'ensemble $U(B)^\circ$, la fonctionnelle $f = U^*(g)$ reste dans la boule unité $B^\circ$ de l'espace $\mathbf{X}^*$, c'est-à-dire

$$U^*(U(B)^\circ) \subset B^\circ \cap U^*(\mathbf{Y}^*).$$

Mettons cette relation sous la forme suivante:

$$U(B)^\circ \subset U^{*-1}(B^\circ \cap U^*(\mathrm{Y}^*)).$$

L'opérateur $U^{*-1}$ étant continu, l'ensemble $S = U^{*-1}(B^\circ \cap \cap U^*(\mathrm{Y}^*))$ est borné dans l'espace $\mathrm{Y}^*$. Donc, le polaire $S^\circ$ est un voisinage de 0 dans l'espace Y. Or

$$\overline{U(B)} = (U(B))^{\circ\circ} \supset S^\circ,$$

c.q.f.d.

Les théorèmes 2 et 2* entraînent le

COROLLAIRE. *Pour que l'opérateur inverse continu $U^{-1}$ existe, il est nécessaire et suffisant qu'existe l'opérateur inverse continu $U^{*-1}$. Et de plus*

$$U^{*-1} = (U^{-1})^*.$$

Prouvons la dernière relation. Soient $f \in \mathrm{X}^*$ et $g = (U^{-1})^*(f)$. On a

$$g(y) = f(U^{-1}(y)) = f(x) \quad (y = U(x) \in \mathrm{Y}).$$

Si, d'autre part, l'on désigne $g' = U^{*-1}(f)$, c'est-à-dire $f = U^*(g')$, alors

$$f(x) = g'(U(x)) = g'(y) \quad (x = U^{-1}(y) \in \mathrm{X}).$$

Pour tout $y \in \mathrm{Y}$ on aura donc $g(y) = g'(y)$, autrement dit $(U^{-1})^*(f) = U^{*-1}(f)$, c.q.f.d.

2.4. Les théorèmes suivants qui établissent les conditions de solubilité des équations (1) et (2) généralisent les théorèmes 2 et 2*.

THÉORÈME 3. *Si l'ensemble $U^*(\mathrm{Y}^*)$ est fermé, alors*

$$U(\mathrm{X}) = \mathrm{N}(\mathrm{N}(U^*)),$$

*c'est-à-dire l'équation* (1) *admet une solution si et seulement si $U^*(g) = 0$ entraîne $g(y) = 0$.*

DÉMONSTRATION. Considérons le sous-espace $\mathrm{Y}_1 = \overline{U(\mathrm{X})}$ de l'espace Y et un opérateur $U_1$ de X dans $\mathrm{Y}_1$:

$$U_1(x) = U(x) \quad (x \in \mathrm{X})\ ^{*)}.$$

Prouvons que

$$U_1^*(\mathrm{Y}_1^*) = U^*(\mathrm{Y}^*). \tag{5}$$

A cet effet désignons par $\omega$ l'opérateur d'immersion de $\mathrm{Y}_1$ dans Y. L'opérateur inverse à gauche continu $\omega^{-1}$ existant de toute évidence, le théorème 2 nous dit que $\omega^*(\mathrm{Y}^*) = \mathrm{Y}_1^*$. Or $U = \omega U_1$, donc

---

*) Les opérateurs $U$ et $U_1$ sont distincts, car les espaces qui contiennent les images $U(\mathrm{X})$ et $U_1(\mathrm{X})$ le sont, de sorte que les opérateurs adjoints $U^*$ et $U_1^*$ opèrent dans des espaces différents.

$U^* = U_1^* \omega^*$ (cf. IX.3.1, tome 1) et par suite

$$U^*(\mathrm{Y}^*) = U_1^*(\omega^*(\mathrm{Y}^*)) = U_1^*(\mathrm{Y}_1^*).$$

Comme $\overline{U_1(\mathrm{X})} = \mathrm{Y}_1$, le théorème 1 nous apprend que $\mathrm{Y}_1 = \mathrm{N}(\mathrm{N}(U_1^*))$, ce qui n'est possible que si $\mathrm{N}(U_1^*) = \{0\}$, donc l'opérateur $U_1^*$ réalise une bijection du $B$-espace $\mathrm{Y}_1^*$ sur l'espace $U_1^*(\mathrm{Y}^*)$, qui, en vertu de (5), est également $B$-espace. Le corollaire du théorème 1.2 nous dit qu'existe l'opérateur inverse à gauche continu $U_1^{*-1}$, donc d'après le théorème 2

$$U(\mathrm{X}) = U_1(\mathrm{X}) = \mathrm{Y}_1.$$

En appliquant le théorème 1 on obtient ce qu'on voulait.

THÉORÈME 3*. *Si l'ensemble $U(\mathrm{X})$ est fermé, alors*

$$U^*(\mathrm{Y}^*) = \mathrm{N}^*(\mathrm{N}(U)), \tag{6}$$

*c'est-à-dire l'équation (2) admet une solution si et seulement si la fonctionnelle $f$ est nulle sur l'ensemble d'annulation de l'opérateur $U$.*

DÉMONSTRATION. Etudions d'abord le cas où $U(\mathrm{X}) = \mathrm{Y}$. Posons $\mathrm{X}_0 = \mathrm{N}(U) = U^{-1}(0)$ et considérons l'espace quotient $\overline{\mathrm{X}} = \mathrm{X}/\mathrm{X}_0$. Construisons, comme indiqué dans 1.3, un opérateur $\overline{U}$ de $\overline{\mathrm{X}}$ sur Y. Ceci étant, $U = \overline{U}\varphi$ où $\varphi$ est un homomorphisme naturel de X sur $\overline{\mathrm{X}}$. L'espace $\overline{\mathrm{X}}$ étant complet et l'opérateur $\overline{U}$ biunivoque, d'après le corollaire du théorème 1.2 l'opérateur inverse continu $\overline{U}^{-1}$ existe et, par suite, en vertu du théorème 2, $\overline{U}^*(\mathrm{Y}^*) = \overline{\mathrm{X}}^*$. Or, $U^* = \varphi^*\overline{U}^*$, donc

$$U^*(\mathrm{Y}^*) = \varphi^*(\overline{U}^*(\mathrm{Y}^*)) = \varphi^*(\overline{\mathrm{X}}^*).$$

Il reste donc à prouver que $\varphi^*(\overline{\mathrm{X}}^*) = \mathrm{N}^*(\mathrm{X}_0)$. Prenons une fonctionnelle arbitraire $f \in \varphi^*(\overline{\mathrm{X}}^*)$. Supposons que $\bar{f} \in \overline{\mathrm{X}}^*$ est telle que $f = \varphi^*(\bar{f})$. Comme

$$f(x) = \bar{f}(\varphi(x)) \quad (x \in \mathrm{X})$$

et que $\varphi(x) = 0$ pour $x \in \mathrm{X}_0$, il vient $f \in \mathrm{N}^*(\mathrm{X}_0)$.

Admettons maintenant que $f$ est une fonctionnelle quelconque de $\mathrm{N}^*(\mathrm{X}_0)$. En reprenant textuellement les raisonnements faits dans 1.3 pour construire l'opérateur $\overline{U}$, on s'assure que la fonctionnelle

$$\bar{f}(\bar{x}) = \bar{f}(x) \quad (x \in \bar{x} = \varphi(x))$$

est définie de façon correcte et est linéaire et continue dans $\overline{\mathrm{X}}$. Mais alors $f = \varphi^*(\bar{f}) \in \varphi^*(\overline{\mathrm{X}}^*)$.

Dans le cas général, si, éventuellement $U(\mathrm{X}) = \mathrm{Y}_1 \neq \mathrm{Y}$, nous introduisons comme dans la démonstration du théorème 3 un opéra-

teur $U_1$ de X sur $Y_1$. L'opérateur $U_1$ entre dans le cas particulier étudié, donc compte tenu de la relation (5), on peut écrire

$$U^*(Y^*) = U_1^*(Y^*) = N^*(N(U_1)) = N^*(N(U)).$$

Ce qui achève la démonstration du théorème.

COROLLAIRE. *Si l'ensemble $U^*(Y^*)$ est fermé, il est (*)-faiblement fermé.*

En effet, le théorème 3 nous dit que l'ensemble $U$ (X) sera fermé, donc on peut se servir du théorème 3* qui affirme que l'ensemble $U^*(Y^*)$ est (*)-faiblement fermé.

Signalons le cas particulier du corollaire où $U^*$ admet un inverse à gauche continu. L'ensemble $U^*(Y^*)$ est fermé dans ce cas aussi, en vertu du corollaire du théorème 1.1.

2.5. Voici quelques corollaires des théorèmes prouvés.

THÉORÈME 4. *Pour que l'équation* (1) *admette une solution unique pour tout $y \in Y$, il est nécessaire et suffisant que l'équation* (2) *possède une solution unique pour tout $f \in X^*$.*

THÉORÈME 5. *Supposons que les équations* (1) *et* (2) *admettent des solutions quels que soient $y \in Y$ et $f \in X^*$. Il existe alors des opérateurs inverses continus $U^{-1}$ et ${U^*}^{-1}$. Donc, les solutions des équations* (1) *et* (2) *sont uniques.*

THÉORÈME 6. *Supposons que les équations* (1) *et* (2) *n'admettent que les solutions triviales pour $y = 0$ et $f = 0$. Si l'un des ensembles $U$(X) ou $U^*(Y^*)$ est fermé, les équations* (1) *et* (2) *admettent des solutions pour tous $y \in Y$ et $f \in X^*$, et ces solutions sont uniques.*

REMARQUE. Dire que l'ensemble $U$ (X) ou $U^*(Y^*)$ est fermé revient à dire (dans les hypothèses du théorème) qu'existe l'opérateur inverse à gauche continu $U^{-1}$ ou ${U^*}^{-1}$.

2.6. Si $X = H$ est un espace hilbertien et $Y = X$, le théorème 6 nous conduit à un résultat obtenu au chapitre IX (voir tome 1), par d'autres raisonnements et même sous une forme plus générale (cf. théorème IX.5.3) : pour qu'un opérateur auto-adjoint $U$ possède un inverse continu, il est nécessaire et suffisant qu'existe l'opérateur inverse à gauche continu $U^{-1}$, c'est-à-dire pour tout $x \in H$ l'on ait

$$\| U(x) \| \geqslant m \| x \| \quad (m > 0).$$

Le théorème 5 entraîne le

THÉORÈME 7. *Si un opérateur auto-adjoint $U$ applique l'espace H sur lui-même, il possède un opérateur inverse continu $U^{-1}$.*

2.7. Appliquons les théorèmes établis dans ce paragraphe à la résolution du système d'équations algébriques linéaires

$$\begin{aligned} a_{11}\xi_1 + a_{12}\xi_2 + \dots + a_{1m}\xi_m &= \eta_1, \\ a_{21}\xi_1 + a_{22}\xi_2 + \dots + a_{2m}\xi_m &= \eta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \dots + a_{nm}\xi_m &= \eta_n. \end{aligned} \quad (7)$$

On peut remplacer ce système par l'équation

$$U(x) = y,$$

où $U$ est un opérateur d'un espace normé $X_m$ de dimension $m$ dans un espace normé $Y_n$ *) de dimension $n$, défini par la matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

L'opérateur adjoint $U^*$ de $Y_n^*$ dans $X_m^*$ est défini par la matrice

$$A^* = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix},$$

de sorte que

$$f = U^*(g) \quad (f = (\varphi_1, \varphi_2, \dots, \varphi_m) \in X_m^*, \quad g = (\psi_1, \psi_2, \dots, \psi_n) \in Y_n^*)$$

signifie que

$$\varphi_k = \sum_{j=1}^{n} a_{jk}\psi_j \quad (k = 1, 2, \dots, m).$$

L'ensemble $U(X_m)$ étant un ensemble linéaire de dimension finie, donc fermé (IV.4.6, tome 1), le théorème 1 donne

$$U(X_m) = N(N(U^*)).$$

Or, $N(U^*)$ est composé des fonctionnelles $g = (\psi_1, \psi_2, \dots, \psi_n) \in Y_n^*$ telles que $U^*(g) = 0$, c'est-à-dire

$$\sum_{j=1}^{n} a_{jk}\psi_j = 0 \quad (k = 1, 2, \dots, m). \quad (8)$$

*) La manière dont la métrique a été introduite dans $X_m$ et $Y_n$ importe peu. Pour fixer les idées on peut admettre que $X_m = l_m^2$, $Y_n = l_n^2$ et que les espaces sont réels.

Donc $y = (\eta_1, \eta_2, \ldots, \eta_n) \in U(\mathrm{X}_m)$ si et seulement si (8) entraîne

$$g(y) = \sum_{j=1}^{n} \psi_j \eta_j = 0.$$

En d'autres termes, pour que le système (7) admette une solution, il est nécessaire et suffisant que ses seconds membres forment un vecteur orthogonal à toute solution du système homogène adjoint.

Supposons maintenant que $m = n$. Si le système (7) admet une solution quels que soient les seconds membres, l'opérateur $U^*$ admet un inverse à gauche (théorème 2*). On voit aussitôt que dans ce cas $U^*(\mathrm{Y}_n^*) = \mathrm{X}_n^*$, sinon $U^*$ appliquerait un espace de dimension $n$ dans un espace de moindre dimension, ce qui contredirait l'existence de l'inverse à gauche $U^{*^{-1}}$ *). En appliquant le théorème 4 on trouve que la solution du système (7) est unique.

En utilisant le théorème 5 on établit la réciproque par des raisonnements analogues: si le système (7) admet une solution au plus, il en admettra une quels que soient les seconds membres.

Pour les systèmes à matrices symétriques ces deux résultats sont une conséquence immédiate du théorème 7.

---

*) En effet, supposons que l'ensemble $U^*(\mathrm{Y}_n^*)$ est de dimension $n_0 < n$, c'est-à-dire il existe des éléments $f_1, f_2, \ldots, f_{n_0} \in U^*(\mathrm{Y}_n^*)$ tels que pour tout $f \in U^*(\mathrm{Y}_n^*)$

$$f = \sum_{k=1}^{n_0} c_k f_k.$$

Pour tout $g \in \mathrm{Y}_n^*$ on a alors $g = \sum_{k=1}^{n_0} c_k g_k$, où $g_k = U^{*^{-1}}(f_k)$, c'est-à-dire $\mathrm{Y}_n^*$ serait de dimension $n$.

HAPITRE XIII

# ÉQUATIONS FONCTIONNELLES DE SECONDE ESPÈCE

Dans ce chapitre nous étudions l'équation

$$x - \lambda U(x) = y, \tag{*}$$

où $U$ est un opérateur linéaire continu d'un $B$-espace X dans lui-même. Nous appelons cette équation, *équation de seconde espèce*, et l'opérateur $U$, *noyau de l'équation*. La définition a été empruntée à la théorie des équations intégrales, où on appelle équation de seconde espèce l'équation

$$x(s) - \lambda \int_a^b K(s, t)\, x(t)\, dt = y(s) \quad (s \in [a, b]),$$

et équation de première espèce, l'équation

$$\int_a^b K(s, t)\, x(t)\, dt = y(s), \quad (s \in [a, b]).$$

Bien que, formellement, l'équation fonctionnelle (*) puisse être ramenée à une équation de « première espèce »

$$T(x) = y \quad (T = I - \lambda U),$$

il apparaît plus rationnel de mettre en évidence l'opérateur identique, car l'opérateur $U$ est susceptible de posséder de meilleures propriétés que l'opérateur $T$, d'où la possibilité d'étudier l'équation (*) de manière plus approfondie.

## § 1. Equations à noyau compact

Dans ce paragraphe nous étudions l'équation

$$x - U(x) = y \quad (x, y \in \mathrm{X}) \tag{1}$$

et son adjointe

$$g - U^*(g) = f \quad (f, g \in \mathrm{X}^*), \tag{2}$$

en admettant que $U$ (donc $U^*$ d'après le théorème IX.3.3, tome 1) est compact dans un $B$-espace X. Posons $T = I - U$, où $I$ représente l'opérateur identique. L'équation (1) prend la forme simple

$$T(x) = y, \tag{1'}$$

et l'équation (2)

$$T^*(g) = f, \tag{2'}$$

puisque $T^* = I^* - U^*$ (IX.3.1, tome 1) et $I^*$ est l'opérateur identique dans $\mathbf{X}^*$.

**1.1.** Nous commençons par démontrer trois lemmes.

Lemme 1. *L'ensemble* $T(\mathbf{X})$ *est fermé.*

Démonstration. Posons $\mathbf{X}_0 = \mathbf{N}(T) = T^{-1}(0)$ et considérons l'espace quotient $\overline{\mathbf{X}} = \mathbf{X}/\mathbf{X}_0$ et l'opérateur $\overline{T}$ de $\overline{\mathbf{X}}$ dans X (cf. XII.1.3). Par $\varphi$ on désigne comme dans XII.1.3 l'homomorphisme naturel de X sur $\overline{\mathbf{X}}$. Soit $\{y_n\} \subset T(\mathbf{X})$ une suite convergeant vers l'élément $y_0 \in \mathbf{X}$. Comme $\overline{T}(\overline{\mathbf{X}}) = T(\mathbf{X})$ il existe des éléments $\overline{x}_n \in \overline{\mathbf{X}}$ tels que $y_n = T(\overline{x}_n)$ $(n = 1, 2, \ldots)$. Trouvons $x_n \in \mathbf{X}$ à partir de la relation (1) de IV.1.8, tome 1, c'est-à-dire tel que

$$\overline{x}_n = \varphi(x_n) \quad \|\overline{x}_n\| \geqslant \frac{1}{2}\|x_n\| \quad (n = 1, 2, \ldots). \tag{3}$$

Prouvons que la suite $\{\overline{x}_n\}$ est bornée. Supposons le contraire. Quitte à passer à une sous-suite, on peut faire en sorte que $c_n = \|\overline{x}_n\| \to \infty$. D'après (3), la suite $\left\{\frac{x_n}{c_n}\right\}$ est bornée, donc en passant encore une fois à une sous-suite on peut admettre que $\left\{U\left(\frac{x_n}{c_n}\right)\right\}$ converge. Supposons par exemple que $U\left(\frac{x_n}{c_n}\right) \to z$. Comme $T(x_n) = \overline{T}(\overline{x}_n) = y_n$ on peut écrire

$$\frac{x_n}{c_n} = U\left(\frac{x_n}{c_n}\right) + T\left(\frac{x_n}{c_n}\right) = U\left(\frac{x_n}{c_n}\right) + \frac{y_n}{c_n} \to z,$$

donc

$$T(z) = \lim_{n\to\infty} T\left(\frac{x_n}{c_n}\right) = \lim_{n\to\infty}\left(\frac{y_n}{c_n}\right) = 0,$$

c'est-à-dire $z \in \mathbf{X}_0$. Mais alors

$$\frac{\overline{x}_n}{c_n} = \varphi\left(\frac{x_n}{c_n}\right) \to \varphi(z) = 0,$$

ce qui est impossible, puisque $\left\|\frac{\overline{x}_n}{c_n}\right\| = 1$, $\forall n = 1, 2, \ldots$

Donc, la suite $\{\overline{x}_n\}$ est bornée et *a fortiori* la suite $\{x_n\}$ en vertu de (3). On peut donc admettre que $\{U(x_n)\}$ est une suite convergente.

Si, par exemple, $U(x_n) \to x$, alors

$$x_n = T(x_n) + U(x_n) = y_n + U(x_n) \to y_0 + x = x_0.$$

D'où il résulte que

$$y_0 = \lim_{n\to\infty} y_n = \lim_{n\to\infty} T(x_n) = T(x_0) \in T(X),$$

c.q.f.d.

LEMME 2. *La suite d'ensembles*

$$N(T),\ N(T^2),\ \ldots,\ N(T^n),\ \ldots$$

*est croissante et ne contient qu'un nombre fini d'ensembles distincts.*

DÉMONSTRATION. La première partie du lemme est presque évidente, car si $x \in N(T^n)$, alors $T^n(x) = 0$ et *a fortiori* $T^{n+1}(x) = 0$, c'est-à-dire $x \in N(T^{n+1})$, $N(T^n) \subset N(T^{n+1})$.

Pour prouver la deuxième partie du lemme, désignons $X_n = N(T^n)$ et voyons si pour un $n = 1, 2, \ldots$ $X_n = X_{n+1}$ entraîne $X_{n+1} = X_{n+2}$. Soit $x \in X_{n+2}$. Cela signifie que $T^{n+2}(x) = T^{n+1}(T(x)) = 0$, donc $T(x) \in X_{n+1} = X_n$. Par suite, $T^{n+1}(x) = T^n(T(x)) = 0$, c'est-à-dire $x \in X_{n+1}$. Donc, $X_{n+2} \subset X_{n+1}$. L'inclusion inverse ayant toujours lieu, on a en définitive $X_n = X_{n+1} = X_{n+2} = \ldots$

Supposons maintenant que pour chaque $n = 1, 2, \ldots$

$$X_n \neq X_{n+1}.$$

Chaque $X_n$ étant un sous-espace de $X_{n+1}$, le lemme de la quasi-perpendiculaire (IV.1.7, tome 1) nous dit que l'on peut exhiber un élément normé $x_{n+1}$ de $X_{n+1}$ tel que

$$\rho(x_{n+1}, X_n) > \frac{1}{2} \quad (n = 0, 1, 2, \ldots). \tag{4}$$

Soit $m > n$. Considérons l'élément

$$U(x_m) - U(x_n) = x_m - T(x_m) - [x_n - T(x_n)] = x_m - \tilde{x},$$

où $\tilde{x} = T(x_m) + x_n - T(x_n)$. Prouvons que $\tilde{x} \in X_{m-1}$. En effet,

$$T^{m-1}(\tilde{x}) = T^m(x_m) + T^{m-1}(x_n) - T^m(x_n) = 0,$$

puisque $x_n \in X_n \subset X_{m-1}$ et $x_m \in X_m$.

Compte tenu de (4) il vient

$$\| U(x_m) - U(x_n) \| = \| x_m - \tilde{x} \| > \frac{1}{2} \quad (m > n;\ m, n = 1, 2, \ldots). \tag{5}$$

Mais $\{x_n\}$ est une suite bornée, donc en vertu de la compacité de l'opérateur $U$, de la suite $\{U(x_n)\}$ on peut extraire une suite partielle convergente, ce qui contredit (5).

LEMME 3. *Parmi les ensembles*

$$T(X),\ T^2(X),\ \ldots,\ T^n(X),\ \ldots \tag{6}$$

*il n'en existe qu'un nombre fini de distincts.*

DÉMONSTRATION. Elle rappelle dans les grands traits celle du lemme précédent, aussi la ferons-nous sans entrer dans les détails.

On remarquera que les ensembles (6) sont fermés d'après le lemme 1 et de plus forment une suite décroissante. Il est clair que l'égalité $T^n(X) = T^{n+1}(X)$ entraîne pour un $n$

$$T^n(X) = T^{n+1}(X) = T^{n+2}(X) = \ldots,$$

et le lemme est prouvé dans ce cas.

Supposons que $T^n(X) \neq T^{n+1}(X)$ $(n = 0, 1, \ldots)$ et utilisons le lemme de la quasi-perpendiculaire (IV.1.7, tome 1) pour construire une suite $\{x_n\}$ telle que

$$\|x_n\| = 1,\quad x_n \in T^n(X),\quad \rho(x_n, T^{n+1}(X)) > \frac{1}{2}\quad (n = 1, 2, \ldots). \tag{7}$$

Soit $m > n$. Comme dans le lemme 2, on a

$$U(x_n) - U(x_m) = x_n - T(x_n) - [x_m - T(x_m)] = x_n - \tilde{x}.$$

Or

$$T(x_n) \in T^{n+1}(X),\quad x_m \in T^m(X) \subset T^{n+1}(X),$$
$$T(x_m) \in T^{m+1}(X) \subset T^{n+1}(X),$$

donc

$$\tilde{x} = T(x_n) + x_m - T(x_m) \in T^{m+1}(X).$$

Il s'ensuit de (7)

$$\|U(x_n) - U(x_m)\| = \|x_n - \tilde{x}\| > \frac{1}{2}\quad (m > n;\ m, n = 1, 2, \ldots),$$

ce qui contredit la compacité de l'opérateur $U$.

1.2. Désignons par $r$ le plus petit des entiers positifs $n$ tels que $T^n(X) = T^{n+1}(X)$. En particulier, si $T(X) = X = T^0(X)$, on pose $r = 0$.

Soit par ailleurs $X' = T^r(X)$, $X'' = N(T^r)$.

Le théorème suivant caractérise l'opérateur $T$, donc l'équation (1).

THÉORÈME 1. a) *L'opérateur $T$ réalise une bijection du sous-espace $X'$ sur lui-même.*

b) *Le sous-espace $X''$ est de dimension finie. L'opérateur $T$ applique $X''$ dans lui-même.*

c) *Chaque élément $x \in X$ peut être mis de façon unique sous la forme*

$$x = x' + x''\quad (x' \in X',\ x'' \in X''); \tag{8}$$

*ceci étant, il existe une constante $M > 0$ telle que*

$$\| x' \| \leqslant M \| x \|, \quad \| x'' \| \leqslant M \| x \|. \tag{9}$$

d) *L'opérateur $U$ admet la représentation*

$$U = U' + U'', \tag{10}$$

*où $U'$ et $U''$ sont des opérateurs compacts, le premier de $\mathrm{X}$ dans $\mathrm{X}'$, le second de $\mathrm{X}$ dans $\mathrm{X}''$. Ceci étant, l'opérateur $T' = I - U'$ possède un inverse continu et*

$$U'U'' = U''U' = 0. \tag{11}$$

DÉMONSTRATION. a) Comme $\mathrm{X}' = T^r(\mathrm{X})$, on a

$$T(\mathrm{X}') = T^{r+1}(\mathrm{X}) = T^r(\mathrm{X}) = \mathrm{X}'.$$

Si $T(x) = 0$, où $x \in \mathrm{X}'$, en prenant $n \geqslant r$ de telle sorte qu'en vertu du lemme 2 l'on ait $\mathrm{N}(T^n) = \mathrm{N}(T^{n+1})$, il vient $x \in T^n(\mathrm{X})$ et, par suite, il existe un $\tilde{x} \in \mathrm{X}$ tel que $x = T^n(\tilde{x})$. Or, $0 = T(x) = T^{n+1}(\tilde{x})$, donc $\tilde{x} \in \mathrm{N}(T^{n+1}) = \mathrm{N}(T^n)$, c'est-à-dire $x = T^n(\tilde{x}) = 0$.

b) On a

$$T^r = (I - U)^r = I - U_1,$$

où l'opérateur $U_1$ est une combinaison linéaire des puissances positives de l'opérateur $U$. Donc, l'opérateur $U_1$ est compact d'après le théorème IX.2.2 (tome 1). Chaque ensemble borné dans $\mathrm{X}''$ est compact, car $U_1(x) = x$ pour $x \in \mathrm{X}''$. D'après le théorème IV.1.3 (tome 1), l'espace $\mathrm{X}''$ est de dimension finie.

Si $r > 0$, l'ensemble $T(\mathrm{X}'')$ est de toute évidence $\mathrm{N}(T^{r-1}) \subset \mathrm{N}(T^r) = \mathrm{X}''$. Si $r = 0$, alors $\mathrm{X}'' = \{0\}$ et l'inclusion $T(\mathrm{X}'') \subset \mathrm{X}''$ est triviale.

c) Désignons par $T_0$ l'opérateur $T$ considéré seulement sur $\mathrm{X}'$. En appliquant le lemme 1 à l'opérateur $T^r = I - U_1$ on déduit que $\mathrm{X}'$ est fermé, donc c'est un $B$-espace. Et, par suite, d'après le théorème XII.1.2, l'opérateur $T_0$ qui réalise une bijection de $\mathrm{X}'$ sur lui-même, admet un inverse continu $T_0^{-1}$.

Soit $x$ un élément quelconque de $\mathrm{X}$; posons

$$x' = T_0^{-r}T^r(x), \quad x'' = x - x' = x - T_0^{-r}T^r(x). \tag{12}$$

Il est clair que $x' \in \mathrm{X}'$ et comme

$$T^r(x'') = T^r(x) - T^rT_0^{-r}T^r(x) = T^r(x) - T^r(x) = 0,$$

alors $x'' \in \mathrm{X}''$, ce qui prouve la représentabilité de $x$ par (8).

Si $x = x_1' + x_1''$ est une autre représentation de $x$ sous la forme (8), $x_1' \in \mathrm{X}'$, $x_1'' \in \mathrm{X}''$, alors

$$T^r(x) = T^r(x_1') + T^r(x_1'') = T^r(x_1').$$

Mais comme $x'_1 \in X'$, alors $T^r(x'_1) = T^r_0(x'_1)$, donc

$$x'_1 = T_0^{-r} T^r(x) = x',$$

ce qui prouve l'unicité de la représentation (8).

Les majorations (9) résultent en vertu de (12), de la continuité de l'opérateur $T_0^{-1}$.

d) Comme $U = I - T$, pour $x \in X'$ on a

$$U(x) = x - T(x) \in X',$$

c'est-à-dire l'opérateur $U$ applique $X'$ dans lui-même. On s'assure de façon analogue que $U(X'') \subset X''$.

Pour un $x$ quelconque de X posons

$$U'(x) = U(x'), \quad U''(x) = U(x''), \tag{13}$$

où $x' \in X'$ et $x'' \in X''$ sont ceux de la représentation de $x$ sous la forme (8). En utilisant les majorations (9), on s'assure sans peine que $U'$ et $U''$ sont des opérateurs linéaires continus. De plus, il est clair que $U = U' + U''$ et $U'(X) \subset X'$, $U''(X) \subset X''$. D'autre part, il est manifeste que

$$U'(X'') = U''(X') = \{0\}. \tag{14}$$

Ces relations impliquent $U'U'' = U''U' = 0$, c'est-à-dire (11).

L'opérateur $U''$ applique l'espace X dans l'espace de dimension finie $X''$ dans lequel tout ensemble borné est compact. Donc, $U''$ est un opérateur compact. Or, $U' = U - U''$, et le théorème IX.2.2 (tome 1) nous dit que l'opérateur $U'$ est compact.

Prouvons enfin que l'opérateur $T' = I - U'$ admet un inverse continu. Pour cela il suffit de montrer, premièrement, que $T'(x) = 0$ entraîne $x = 0$ et, deuxièmement, que $T'(X) = X$. Supposons que $T'(x) = 0$. En mettant $x$ sous la forme (8), on obtient

$$0 = T'(x) = x - U'(x) = x' - U(x') + x'' = T(x') + x''.$$

Comme $T(x') \in X'$, l'unicité de la représentation de l'élément 0 sous la forme (8) donne

$$T(x') = x'' = 0$$

et en vertu du point a) $x' = 0$. Donc $x = x' + x'' = 0$.

Soit maintenant un élément $y$ quelconque de X. Mettons-le sous la forme (8), soit $y = y' + y''$ ($y' \in X'$, $y'' \in X''$) et posons

$$x = T_0^{-1}(y') + y''.$$

Comme $T_0^{-1}(y') \in X'$, il vient

$$U'(x) = U(T_0^{-1}(y'))$$

et

$$T'(x) = x - U'(x) = T_0^{-1}(y') - U(T_0^{-1}(y')) + y'' =$$
$$= TT_0^{-1}(y') + y'' = y' + y'' = y.$$

Donc, $T'(\mathrm{X}) = \mathrm{X}$. Ce qui achève la démonstration du théorème.

Remarque. Supposons que $m$ est le plus petit des entiers positifs $n$ tels que $\mathrm{N}(T^n) = \mathrm{N}(T^{n+1})$. Alors $m = r$.

En effet, en prenant $x \in \mathrm{N}(T^{r+1})$ et en le mettant sous la forme (8), on obtient

$$0 = T^{r+1}(x) = T^{r+1}(x') + T^{r+1}(x'') = T^{r+1}(x'),$$

ce qui, en vertu de a), n'est possible que pour $x' = 0$. D'où $x = x'' \in \in \mathrm{N}(T^r)$ et par suite $m \leqslant r$.

Si $y = T^m(x)$ $(x \in \mathrm{X})$, en remplaçant $x$ par sa forme (8) on a

$$y = T^m(x) = T^m(x') + T^m(x'') = T^m(x') = T^{m+1}(T_0^{-1}(x'))$$

et $y \in T^{m+1}(\mathrm{X})$, donc on doit avoir $r \leqslant m$.

Le cas particulier du fait indiqué dans la remarque est exprimé dans le théorème suivant.

Théorème 2. *Pour que l'équation* (1) *admette une solution quel que soit* $y \in \mathrm{X}$ *il est nécessaire et suffisant que l'équation homogène*

$$T(x) = 0 \qquad (15)$$

*admette une solution unique (de toute évidence* $x = 0$*).*

En effet, si l'équation (1) admet une solution pour tout $y \in \mathrm{X}$, c'est que $T(\mathrm{X}) = \mathrm{X}$, c'est-à-dire $r = 0$. L'unicité de la solution de l'équation (15) équivaut à $m = 0$.

Remarque. En se servant des résultats du § 2 du chapitre précédent on peut prouver ce théorème, indépendamment du théorème 1, en faisant intervenir le seul fait que l'ensemble $T(\mathrm{X})$ est fermé. Nous laissons au lecteur le soin de faire cette démonstration.

**1.3.** Le théorème suivant établit un lien entre les équations (1) et (2).

Théorème 3. *Les ensembles* $\mathrm{N}(T)$ *et* $\mathrm{N}(T^*)$ *sont de même dimension finie.*

Démonstration. $\mathrm{N}(T)$ est de dimension finie, car $\mathrm{N}(T) \subset \mathrm{N}(T^r) = \mathrm{X}''$ et $\mathrm{X}''$ est de dimension finie en vertu de b) du théorème 1. L'opérateur $U^*$ étant compact, ce qu'on a dit vaut pour l'ensemble $\mathrm{N}(T^*)$.

Supposons que l'ensemble $\mathrm{N}(T)$ est de dimension $n$ et l'ensemble $\mathrm{N}(T^*)$ de dimension $m$. Soient $x_1, x_2, \ldots, x_n$ un système d'éléments linéairement indépendants de $\mathrm{N}(T)$, $g_1, g_2, \ldots, g_m$ un système d'éléments linéairement indépendants de $\mathrm{N}(T^*)$.

Les éléments $x_1, x_2, \ldots, x_n$ étant linéairement indépendants, ils sont justiciables du théorème de biorthogonalisation (théorème V.7.4, tome 1) qui affirme l'existence d'un système biorthogonal de fonctionnelles $f_1, f_2, \ldots, f_n$:

$$f_j(x_k) = \begin{cases} 1, & j=k, \\ 0, & j \neq k \end{cases} \quad (j, k = 1, 2, \ldots, n). \tag{16}$$

De façon analogue, en utilisant le lemme III.3.1, tome 1, on trouve des éléments $y_1, y_2, \ldots, y_m$, tels que

$$g_j(y_k) = \begin{cases} 1, & j=k, \\ 0, & j \neq k \end{cases} \quad (j, k = 1, 2, \ldots, m). \tag{17}$$

Supposons tout d'abord que $n < m$. Soit dans l'espace X l'opérateur $V = U + W$, où

$$W(x) = \sum_{k=1}^{n} f_k(x)\, y_k \quad (x \in \mathrm{X}).$$

L'opérateur linéaire $W$ est compact, car il applique X dans un espace de dimension finie. Donc, $V$ est compact. Considérons l'équation

$$\widetilde{T}(x) = x - V(x) = T(x) - \sum_{k=1}^{n} f_k(x)\, y_k = 0. \tag{18}$$

Soit $x_0$ une solution:

$$\widetilde{T}(x_0) = T(x_0) - \sum_{k=1}^{n} f_k(x_0)\, y_k = 0. \tag{19}$$

Cette égalité entraîne

$$g_s(T(x_0)) - \sum_{k=1}^{n} f_k(x_0)\, g_s(y_k) = 0 \quad (s = 1, 2, \ldots, n), \tag{20}$$

c'est-à-dire, eu égard à (17),

$$g_s(T(x_0)) - f_s(x_0) = 0.$$

d'où, puisque $T^*(g_s) = 0$,

$$f_s(x_0) = 0 \quad (s = 1, 2, \ldots, n). \tag{21}$$

Ceci et (19) donnent $T(x_0) = 0$. c'est-à-dire $x_0 \in \mathrm{N}(T)$, donc $x_0$ peut être représenté sous la forme

$$x_0 = \sum_{h=1}^{n} \alpha_h x_h.$$

Comme $\alpha_s = f_s(x_0)$ d'après (16), il résulte de (21) que $\alpha_s = 0$, donc $x_0 = 0$. Par suite, l'équation (18) admet une solution unique. D'après le théorème 2, l'équation non homogène associée admet une

solution quel que soit le second membre. En particulier, l'équation

$$\tilde{T}(x) = T(x) - \sum_{k=1}^{n} f_k(x)\, y_k = y_{n+1}$$

admet une solution. Appelons-la $x^*$. D'une part

$$g_{n+1}\Big(T(x^*) - \sum_{k=1}^{n} f_k(x^*)\, y_k\Big) = T^*(g_{n+1})(x^*) - \\ - \sum_{k=1}^{n} f_k(x^*)\, g_{n+1}(y_k) = 0,$$

de l'autre

$$g_{n+1}(y_{n+1}) = 1.$$

Donc, on doit avoir $m \leqslant n$.

Par des raisonnements analogues, on exclut le cas $m < n$. Plus exactement, au lieu de l'équation (13) il faut considérer l'équation

$$T^*(g) - \sum_{k=1}^{m} g(y_k)\, f_k = 0$$

dans l'espace $X^*$.

1.4. En groupant les théorèmes établis ci-dessus on obtient le

THÉORÈME 4. *Ou bien les équations* (1) *et* (2) *admettent des solutions pour tout second membre et ces solutions sont uniques. Ou bien les équations homogènes*

$$T(x) = 0 \quad et \quad T^*(g) = 0$$

*possèdent le même nombre fini de solutions linéairement indépendantes* $x_1, x_2, \ldots, x_n$ *et* $g_1, g_2, \ldots, g_n$ *respectivement. Ceci étant, pour que l'équation* (1) (*resp. l'équation* (2)) *admette une solution, il est nécessaire et suffisant que*

$$g_k(y) = 0 \quad (k = 1, 2, \ldots, n)$$

(*respectivement*

$$f(x_k) = 0 \quad (k = 1, 2, \ldots, n)).$$

*Ceci étant, la solution générale de l'équation* (1) *est*

$$x = x^* + \sum_{k=1}^{n} c_k x_k,$$

*celle de l'équation* (2)

$$g = g^* + \sum_{k=1}^{n} c_k g_k,$$

*où* $x^*$ (*resp.* $g^*$) *est une solution quelconque de l'équation* (1) (*resp.* (2)) *et* $c_1, c_2, \ldots, c_n$ *des constantes arbitraires.*

On démontre la deuxième partie du théorème en appliquant aux équations (1) et (2) les théorèmes XII.2.3 et XII.2.3* dont les hypothèses sont remplies en vertu du lemme 1 (mais on peut aussi bien la prouver directement).

Ce théorème est appelé *alternative de Fredholm* pour son analogie avec le théorème classique de la théorie des équations intégrales.

## § 2. Sur les espaces normés complexes

Pour des considérations qui apparaîtront ultérieurement (cf. § 3) il semble naturel d'étudier l'équation

$$x - \lambda U(x) = y$$

dans un espace complexe en attribuant notamment des valeurs complexes à $\lambda$. A ce propos, nous introduisons plus bas quelques notions complémentaires sur les espaces complexes qui nous permettront d'inclure le cas réel dans le complexe.

**2.1.** Soit Z un espace normé complexe. Nous dirons que Z possède un *noyau réel* si sur Z est défini un opérateur $C$ de Z dans lui-même, appelé *involution*, et possédant les propriétés suivantes:

1. $C(\lambda_1 z_1 + \lambda_2 z_2) = \bar{\lambda}_1 C(z_1) + \bar{\lambda}_2 C(z_2) \quad (z_1, z_2 \in Z)$.
2. $C^2(z) = z \quad (z \in Z)$.
3. $\| C(z) \| = \| z \|$.

L'ensemble de tous les éléments pour lesquels $C(z) = z$ s'appelle *noyau réel* de l'espace Z et se note Re Z; les éléments de cet ensemble sont dits *réels*.

Soit $z \in Z$; l'élément $x = \frac{1}{2}[z + C(z)]$ est appelé *partie réelle* de $z$ et noté Re $z$. L'élément $y = \frac{1}{2i}[z - C(z)]$ *partie imaginaire* de $z$ et noté $y = \mathrm{Im}\, z$. Comme

$$C(x) = \frac{1}{2}[C(z) + C^2(z)] = x, \quad C(y) = -\frac{1}{2i}[C(z) - C^2(z)] = y,$$

alors $x = \mathrm{Re}\, z$ et $y = \mathrm{Im}\, z$ sont réels. De toute évidence

$$z = x + yi \tag{1}$$

et

$$C(z) = x - yi.$$

La dernière relation exprime que $C(z)$ est *conjugué* complexe de $z$, c'est-à-dire $C(z) = \bar{z}$.

Si un élément $z$ est représenté sous la forme (1) avec $x$ et $y$ réels, on a automatiquement $x = \mathrm{Re}\, z$ et $y = \mathrm{Im}\, z$. En effet, $\bar{z} = \bar{x} - \overline{yi} =$

$= x - yi$, donc $x = \frac{1}{2}(z + \bar{z}) = \operatorname{Re} z$, $y = \frac{1}{2i}(z - \bar{z}) = \operatorname{Im} z$. Donc, la représentation (1) est unique ; les éléments $x$ et $y$ sont définis de façon unique par $z$.

Le noyau réel X de l'espace Z est un espace normé réel complet si l'espace initial Z est complet.

En effet, toute combinaison linéaire d'éléments de X à coefficients réels est un élément de X. La réalisation des axiomes d'un espace normé pour X résulte de ce que ces axiomes sont satisfaits pour l'espace Z. Assurons-nous que X est complet (sous l'hypothèse de la complétude de Z). Soit $\{x_n\}$ une suite de Cauchy d'éléments de X. En la considérant dans Z complet, on trouve qu'il existe $\lim_{n\to\infty} x_n = z \in Z$. Comme $\overline{x_n - z} = \bar{x}_n - \bar{z}$ d'après la condition 1, il suit que $\| x_n - z \| = \| \overline{x_n - z} \| = \| x_n - \bar{z} \|$, d'où $x_n \to \bar{z}$. La limite étant unique, on a $\bar{z} = z$, c'est-à-dire $z \in X$.

Tous les espaces réels envisagés plus haut sont les noyaux réels d'espaces complexes. Ainsi, C $(K)$ réel est le noyau réel de l'espace complexe $C_{\mathbb{C}}(K)$, l'involution étant l'application $x(t) \to \overline{x(t)}$. On peut en dire autant des espaces $L^p$, V, $c_0$, etc.

D'une façon générale, tout espace réel X peut être considéré comme noyau réel d'un espace complexe, plus exactement de l'espace Z dont les éléments sont les couples ordonnées d'éléments de X : $z = (x, y)$ $(x, y \in X)$, cet espace étant muni des lois de composition suivantes : $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$, $\lambda(x, y) = (\alpha x - \beta y, \beta x + \alpha y)$ $(\lambda = \alpha + \beta i$, $\alpha$, $\beta$ sont réels) et de plus

$$\| (x, y) \| = \max_{\theta} \| x \cos \theta + y \sin \theta \| \text{ *)}. \tag{2}$$

Pour s'assurer que X est un noyau réel de Z il suffit de poser

$$C((x, y)) = (x, -y).$$

Ceci étant, les éléments de la forme $(x, 0)$ et, eux seuls, seront réels dans l'espace Z. Comme

$$\alpha(x_1, 0) + \beta(x_2, 0) = (\alpha x_1 + \beta x_2, 0),$$
$$\| (x, 0) \| = \max_{\theta} \| x \cos \theta \| = \| x \|,$$

en identifiant $(x, 0)$ et $x$, on obtient ce qu'on voulait. Une telle identification permet d'utiliser $x + yi$ au lieu de $(x, y)$. L'espace Z est appelé *complexifié* de X.

*) Il est bien entendu que la norme peut être introduite d'une manière différente dans les autres espaces. Si, par exemple, $X = L^p$, alors la norme usuelle dans $L^p_{\mathbb{C}}$ complexe est équivalente mais non égale à celle définie par (2).

**2.2.** Soient **Z** et **W** des espaces complexes à noyaux réels **X** = = Re **Z** et **Y** = Re **W**. Un opérateur linéaire continu $\widetilde{U}$ de **Z** dans **W** est par définition réel s'il applique les éléments réels de **Z** dans ceux de **W**, c'est-à-dire si $\widetilde{U}$ (**X**) $\subset$ **Y**; un opérateur réel induit donc un opérateur linéaire continu de l'espace réel **X** dans l'espace réel **Y**.

Inversement, si $U$ est un opérateur linéaire continu de **X** dans **Y** et si **X** = Re **Z** et **Y** = Re **W**, alors en posant

$$\widetilde{U}(z)=\widetilde{U}(x+yi)=U(x)+iU(y) \quad (z=x+yi),$$

on obtient un opérateur linéaire continu $\widetilde{U}$ de l'espace complexe **Z** dans l'espace complexe **W**. Les opérateurs $\widetilde{U}$ et $U$ sont visiblement confondus sur **X**. L'opérateur $\widetilde{U}$ est dit *prolongement complexe* de $U$.

Signalons l'inégalité

$$\| U \| \leqslant \| \widetilde{U} \| \leqslant 2 \| U \|. \tag{3}$$

La première partie de l'inégalité est évidente. La deuxième est une conséquence de la suite d'inégalités

$$\| \widetilde{U}(z) \| = \| U(x)+iU(y) \| \leqslant \| U(x) \| + \\ + \| U(y) \| \leqslant \| U \| (\| x \| + \| y \|) \leqslant 2 \| U \| \, \| z \|.$$

(Nous avons utilisé l'inégalité $\| \mathrm{Re}\, z \| \leqslant \| z \|$ qui se démontre comme suit : $\| \mathrm{Re}\, z \| = \frac{1}{2} \| z+\bar{z} \| \leqslant \frac{1}{2} [\, \|z\| + \|\bar{z}\| \,] = \| z \|$.)

Dans nombre de cas on peut démontrer que $\| \widetilde{U} \| = \| U \|$. Si la norme dans les espaces **Z** et **W** est définie par la formule (2), alors

$$\| \widetilde{U}(z) \| = \max_{\theta} \| U(x) \cos\theta + U(y) \sin\theta \| \leqslant \\ \leqslant \| U \| \max_{\theta} \| x \cos\theta + y \sin\theta \| = \| U \| \, \| z \|,$$

donc $\| \widetilde{U} \| \leqslant \| U \|$, ce qui avec (3) donne $\| \widetilde{U} \| = \| U \|$.

Si pour tout $\varepsilon > 0$ il existe un élément normé réel $x$, tel que $\| \widetilde{U}(x) \| > \| \widetilde{U} \| - \varepsilon$, alors $\| \widetilde{U} \| = \| U \|$, puisque dans ce cas

$$\| \widetilde{U} \| \leqslant \| \widetilde{U}(x) \| + \varepsilon = \| U(x) \| + \varepsilon \leqslant \| U \| + \varepsilon.$$

Ceci est réalisé si, par exemple, **Z** et **W** sont des espaces fonctionnels du nombre de ceux indiqués plus haut et $\widetilde{U}$ un opérateur intégral

$$w=\widetilde{U}(z), \qquad w(s)=\int_a^b K(s,\, t)\, x(t)\, dt$$

de noyau réel $K(s,\, t)$.

Cette condition est également réalisée si **Z** et **W** sont des espaces de suites et $\tilde{U}$ un opérateur défini par une matrice réelle.

**2.3.** Si un opérateur réel $\tilde{U}$ est compact, il est évident que l'opérateur $U$ de **X** = Re **Z** dans **Y** = Re **W** qu'il induit est également compact. La réciproque est vraie, c'est-à-dire si $U$ est un opérateur compact de **X** dans **Y**, son prolongement complexe, l'opérateur $\tilde{U}$, est compact.

Pour prouver ce fait il suffit de se rappeler que la convergence de la suite $\{z_n\}$ résulte de celle des suites $\{\mathrm{Re}\, z_n\}$ et $\{\mathrm{Im}\, z_n\}$ et inversement.

**2.4.** Si **Z** est un espace à noyau réel, l'espace dual **Z*** sera aussi à noyau réel.

En effet, soit $f \in \mathbf{Z}^*$ ; définissons l'involution en posant $C(f) = \bar{f}$:

$$\bar{f}(z) = \overline{f(\bar{z})} \quad (z \in \mathbf{Z}). \tag{4}$$

Assurons-nous tout d'abord que $\bar{f}$ est une fonctionnelle linéaire continue. En effet

$$\bar{f}(\lambda z_1 + \mu z_2) = \overline{f(\overline{\lambda z_1 + \mu z_2})} = \overline{f(\bar{\lambda}\bar{z}_1 + \bar{\mu}\bar{z}_2)} =$$
$$= \overline{\bar{\lambda} f(\bar{z}_1) + \bar{\mu} f(\bar{z}_2)} = \lambda \bar{f}(z_1) + \mu \bar{f}(z_2)$$

et

$$|\bar{f}(z)| = |f(\bar{z})| \leqslant \|f\| \|\bar{z}\| = \|f\| \|z\|. \tag{5}$$

Vérifions maintenant que les conditions 1 à 3 de la définition de l'involution sont satisfaites. Compte tenu de la règle de multiplication d'une fonctionnelle par un nombre complexe, on a

$$\overline{(\lambda f_1 + \mu f_2)}(z) = \overline{(\lambda f_1 + \mu f_2)(\bar{z})} = \overline{\lambda f_1(\bar{z}) + \mu f_2(\bar{z})} =$$
$$= \bar{\lambda}\overline{f_1(\bar{z})} + \bar{\mu}\overline{f_2(\bar{z})} = \bar{\lambda}\bar{f}_1(z) + \bar{\mu}\bar{f}_2(z) = (\bar{\lambda}\bar{f}_1 + \bar{\mu}\bar{f}_2)(z),$$
$$\bar{\bar{f}}(z) = \overline{\bar{f}(\bar{z})} = \overline{\overline{f(\bar{\bar{z}})}} = f(z).$$

De (5) enfin il résulte que $\|\bar{f}\| \leqslant \|f\|$. D'autre part $\|f\| = \|\bar{\bar{f}}\| \leqslant \|\bar{f}\|$, d'où $\|\bar{f}\| = \|f\|$.

Désignons par **X** le noyau réel de l'espace **Z** et prouvons que, si $\varphi$ est une fonctionnelle linéaire dans **X**, le prolongement complexe $f$ de $\varphi$ appartient au noyau réel de **Z*** qui ne contient que des fonctionnelles de la forme indiquée. Ces deux propositions découlent immédiatement de la définition d'une fonctionnelle conjuguée. En effet, si $f$ est le prolongement complexe de $\varphi \in \mathbf{X}^*$, on a

$$\bar{f}(z) = \overline{f(\bar{z})} = \overline{\varphi(x) - \varphi(y)\, i} = \varphi(x) + \varphi(y)\, i = f(z)$$
$$(z = x + yi \in \mathbf{Z}).$$

Au contraire, si la fonctionnelle $f$ est un élément réel de l'espace $\mathbf{Z}^*$, c'est-à-dire si $\bar{f} = f$, c'est qu'en vertu de (4) $\overline{f(\bar{z})} = f(z)$. En particulier, si $z = x \in \mathbf{X}$, alors $\overline{f(x)} = f(x)$, c'est-à-dire l'élément $f(x)$ est réel.

La proposition prouvée peut être énoncée sous une autre forme: le noyau réel de l'espace dual est dual du noyau réel de l'espace donné.

**2.5.** L'opérateur adjoint d'un opérateur réel sera naturellement réel. Soient $\tilde{U}$ un opérateur réel de $\mathbf{Z}$ dans $\mathbf{W}$, $U$ l'opérateur de $\mathbf{X}$ dans $\mathbf{Y}$ ($\mathbf{X} = \operatorname{Re} \mathbf{Z}$, $\mathbf{Y} = \operatorname{Re} \mathbf{W}$) qu'il induit. Prouvons que $\tilde{U}^*$ est un opérateur réel et que l'opérateur de $\mathbf{Y}^*$ dans $\mathbf{X}^*$ qu'il induit n'est autre que $U^*$. La dernière proposition doit être comprise au sens que si $\varphi$ et $\psi$ sont respectivement des fonctionnelles dans $\mathbf{X}$ et dans $\mathbf{Y}$ telles que

$$\varphi = U^*(\psi), \tag{6}$$

alors

$$f = \tilde{U}^*(g), \tag{7}$$

où $f$ est le prolongement complexe de $\varphi$, et $g$, celui de $\psi$. Inversement si les fonctionnelles réelles $f \in \mathbf{Z}^*$ et $g \in \mathbf{W}^*$ sont reliées entre elles par la relation (7), les fonctionnelles $\varphi \in \mathbf{X}^*$ et $\psi \in \mathbf{Y}^*$ induites par elles sont reliées par la relation (6).

La preuve que $\tilde{U}^*$ est réel est élémentaire: si $g \in \mathbf{W}^*$ est une fonctionnelle réelle, c'est-à-dire si $\bar{g} = g$ et $f = \tilde{U}^*(g)$, alors

$$\bar{f}(z) = \overline{f(\bar{z})} = \overline{g(\tilde{U}(\bar{z}))} = g(\overline{\tilde{U}(\bar{z})}) = g(\tilde{U}(z)) = f(z).$$

Supposons par ailleurs que $\varphi = U^*\psi$; on a alors ($x = \operatorname{Re} z$, $y = \operatorname{Im} z$)

$$f(z) = \varphi(x) + i\varphi(y) = \psi(U(x)) + i\psi(U(y)) = \\ = g(U(x) + iU(y)) = g(\tilde{U}(z)),$$

de sorte que $f = \tilde{U}^*(g)$. On démontre de façon aussi élémentaire que (7) entraîne (6).

## § 3. Spectre

**3.1.** Dans ce paragraphe et dans le suivant nous étudions le comportement de l'équation

$$T_\lambda(x) = x - \lambda U(x) = y, \tag{1}$$

ou ce qui revient au même, de l'équation

$$T'_\mu(x) = \mu x - U(x) = y \tag{1'}$$

en fonction du paramètre complexe $\lambda$ (ou $\mu$). Ici et dans la suite, $U$ représente un opérateur linéaire continu dans un $B$-espace complexe

X *). Nous considérons les équations (1) et (1′), car il est d'usage d'étudier la première en théorie des équations intégrales, dont nous donnerons quelques applications au § 6, et la deuxième en analyse fonctionnelle abstraite lors de l'examen des propriétés spectrales de l'opérateur $U$.

S'agissant de la solubilité de l'équation (1), on partage le plan complexe en deux ensembles: l'ensemble $\pi(U)$ des valeurs de $\lambda$ pour lesquelles l'équation (1) possède une solution unique quel que soit le second membre $y \in X$ (donc l'opérateur $T_\lambda$ possède un inverse continu (cf. XII.1.3)), et l'ensemble $\chi(U)$ des autres valeurs de $\lambda$. Les points de $\pi(U)$ sont appelés *valeurs non singulières* de $U$; l'ensemble $\chi(U)$ est appelé *ensemble caractéristique* de $U$.

On considère de façon analogue l'ensemble $\rho(U)$ des $\mu$ pour lesquels l'équation (1′) admet une solution unique quel que soit le second membre, et l'ensemble complémentaire $\sigma(U)$. Les points de l'ensemble $\rho(U)$ sont dits *valeurs régulières* de $U$ et l'ensemble $\rho(U)$ *résolvante* de $U$; l'ensemble $\sigma(U)$ est le *spectre* de $U$.

Si pour un $\lambda$ l'équation homogène

$$T_\lambda(x) = x - \lambda U(x) = 0 \tag{2}$$

possède des solutions non nulles, on dit que $\lambda$ est *valeur caractéristique* de $U$. Il est évident que l'ensemble $\chi_0(U)$ des valeurs caractéristiques est contenu dans l'ensemble $\chi(U)$. Chaque solution de l'équation (2) est dite *élément propre*, ou *vecteur propre*, associé à la valeur caractéristique $\lambda$. L'ensemble $\bigcup_{n=1}^{\infty} \mathrm{N}(T_\lambda^n)$ s'appelle *sous-espace radical*, et sa dimension (finie ou infinie) *multiplicité* de la valeur caractéristique $\lambda$. Le nombre $r$ d'ensembles $\mathrm{N}(T_\lambda^n)$ ($n = 1, 2, \ldots$) distincts s'appelle *rang* de la valeur caractéristique $\lambda$ (cf. lemme 2 de XIII.1.1).

Si l'on considère l'équation homogène associée à (1′):

$$T'_\mu(x) = \mu x - U(x) = 0, \tag{2′}$$

on arrive à la notion de *valeur propre*, *d'élément* (*de vecteur*) *propre* et de *sous-espace radical* associés à une valeur propre, qui ont été déjà définis pour des opérateurs dans un espace hilbertien (IX.4.1, tome 1).

A noter que si $U$ est un opérateur auto-adjoint dans un espace hilbertien et $\mu$ sa valeur propre, alors son rang $r = 1$, c'est-à-dire

$$\mathrm{N}(T'_\mu) = \mathrm{N}(T'^2_\mu) = \ldots = \mathrm{N}(T'^n_\mu) = \ldots \tag{3}$$

et par conséquent dans ce cas le sous-espace radical est $\mathrm{N}(T'_\mu)$, c'est-à-dire est constitué de tous les éléments propres de l'opérateur $U$.

Prouvons la relation (3). Les valeurs propres d'un opérateur auto-adjoint étant réelles, l'opérateur $T'_\mu$ et ses puissances seront des opé-

*) Si X est réel au départ, au lieu de $U$ on considère son prolongement complexe.

rateurs auto-adjoints. En supprimant l'indice $\mu$ pour alléger l'écriture et en prenant $n = 2^k$ on aura pour $x \in \mathrm{N}\,(T'^{2^k})$

$$(T'^{2^{k-1}}x,\ T'^{2^{k-1}}x) = (T'^{2^k}x,\ x) = 0,$$

d'où

$$T'^{2^{k-1}}x = 0.$$

En poursuivant cette procédure, on arrive finalement à l'égalité $T'x = 0$, c'est-à-dire $x \in \mathrm{N}\,(T')$, donc $\mathrm{N}\,(T') \supset \mathrm{N}\,(T'^{2^k})$. L'inclusion inverse est réalisée pour un opérateur arbitraire. Donc, $\mathrm{N}\,(T') = \mathrm{N}\,(T'^{2^k})$ $(k = 1, 2, \ldots)$. Si ensuite on prend $n = 2, 3, \ldots$ quelconque et que l'on choisisse $k$ tel que $n \leqslant 2^k$, on aura alors l'inclusion évidente $\mathrm{N}\,(T') \subset \mathrm{N}\,(T'^n) \subset \mathrm{N}\,(T'^{2^k})$. D'où $\mathrm{N}\,(T') = \mathrm{N}\,(T'^n)$.

Indiquons une relation simple entre le spectre et l'ensemble caractéristique d'un opérateur $U$. Il est aisé de voir que si $\lambda \in \chi\,(U)$, alors $\mu = \frac{1}{\lambda} \in \sigma\,(U)$ et inversement. De toute évidence les valeurs propres et les valeurs caractéristiques de l'opérateur $U$ sont reliées par une relation analogue. Ceci étant, il importe d'avoir à l'esprit que l'élément propre associé à une valeur caractéristique $\lambda$ sera également élément propre associé à la valeur propre $\mu = 1/\lambda$ et inversement. Cette remarque est également valable pour les sous-espaces propres, car pour $\lambda\mu = 1$ on a $\mathrm{N}\,(T_\lambda^n) = \mathrm{N}\,(T'^n_\mu)$ $(n = 1, 2, \ldots)$. Il ne sera donc pas nécessaire de faire une distinction entre les notions d'élément propre (caractéristique) associé à une valeur caractéristique et d'élément propre associé à une valeur propre, ce que du reste l'on a fait dans la terminologie introduite plus haut.

La relation indiquée entre l'ensemble caractéristique et le spectre permet de considérer, en fonction de la commodité, l'une seulement de ces notions parallèles mais en fait équivalentes, et de ne donner qu'exceptionnellement les deux formulations.

Indiquons quelques propositions simples relatives aux notions introduites plus haut.

I. La relation $\lambda \in \pi\,(U)$ équivaut à l'existence de l'opérateur inverse continu: $B_\lambda = T_\lambda^{-1} = (1 - \lambda U)^{-1}$.

Cf. XII.1.3, corollaire du théorème XII.1.2.

II. L'ensemble des valeurs non singulières est ouvert, donc l'ensemble caractéristique est fermé.

Ceci découle du théorème qui dit que si un opérateur possède un inverse continu, tout opérateur voisin en norme admet également un inverse continu (V.4.6, tome 1). Ici

$$\| T_\lambda - T_{\lambda'} \| = | \lambda - \lambda' |\, \| U \|,$$

donc si $T_\lambda^{-1}$ existe, $T_{\lambda'}^{-1}$ existera aussi pour une différence $\lambda - \lambda'$ assez petite.

III. Le disque $|\lambda| < \frac{1}{\|U\|}$ est contenu dans l'ensemble $\pi(U)$ ; donc le spectre $\sigma(U)$ est entièrement compris dans le disque $|\mu| \leqslant \|U\|$.

Pour établir cette proposition il suffit d'appliquer le théorème de Banach de l'opérateur inverse (V.4.5, tome 1).

IV. Les ensembles $\chi(U)$ et $\chi(U^*)$ sont symétriques par rapport à l'axe réel.

En effet (IX.3.1, tome 1)

$$T_{\lambda}^* = [I - \lambda U]^* = I^* - \bar{\lambda} U^*.$$

Or, d'après le théorème XII.2.4, les opérateurs $T_{\bar{\lambda}}^{-1}$ et $(T_{\lambda}^*)^{-1}$ *) existent simultanément.

V. Si X est un espace à noyau réel et $U$ un opérateur réel, l'ensemble $\chi_0(U)$ est symétrique par rapport à l'axe réel. De plus, si $\lambda \in \chi_0(U)$ et $z$ est l'élément propre associé, alors à la valeur caractéristique $\bar{\lambda}$ sera associé l'élément propre $\bar{z}$.

En effet,

$$\bar{z} - \bar{\lambda} U(\bar{z}) = \overline{z - \lambda U(z)},$$

d'où il résulte que les égalités $z - \lambda U(z) = 0$ et $\bar{z} - \bar{\lambda} U(\bar{z}) = 0$ sont équivalentes.

Remarque. Au IX.5.3, tome 1, on a donné la définition du spectre pour un opérateur auto-adjoint dans un espace hilbertien. Le théorème IX.5.3 exprime l'équivalence des deux définitions (celle du chapitre IX, tome 1, et la présente).

**3.2.** Si $U$ est un opérateur compact, on peut étudier la structure de l'ensemble caractéristique d'une manière assez complète.

Théorème 1. *Si U est un opérateur compact, alors*

a) *l'ensemble caractéristique n'est composé que de valeurs caractéristiques, c'est-à-dire* $\chi(U) = \chi_0(U)$ ; *de plus, toute valeur caractéristique est de multiplicité finie*;

b) *le disque* $|\lambda| \leqslant r$, $\forall r \neq 0$, *ne contient qu'un nombre fini de valeurs caractéristiques*;

c) *si* $\lambda_1 \in \chi(U)$ *et* $\lambda_2 \in \chi(U^*)$, $\lambda_1 \neq \bar{\lambda}_2$, *et si* $x_1 \in$ X *est l'élément propre associé à* $\lambda_1$, $g_2 \in$ X* *l'élément propre associé à* $\lambda_2$, *alors*

$$g_2(x_1) = 0.$$

Démonstration. a) Le théorème XII.1.4 nous dit que si $\lambda \notin \pi(U)$, l'équation homogène (2) possède une solution non nulle. Que le sous-espace propre soit de dimension finie découle du lemme 2 (XIII.1.1). En effet, d'après ce lemme il existe un $n_0$ tel que $\mathrm{N}(T_{\lambda}^n) =$

*) $T_{\lambda}^*$ désigne $(T_{\lambda})^*$.

$= \mathrm{N}(T_\lambda^{n_0})$ $(n \geqslant n_0)$. Donc, le sous-espace propre est ici $\mathrm{N}(T_\lambda^{n_0})$. Or

$$T_\lambda^{n_0} = (I - \lambda U)^{n_0} = I - \tilde{U},$$

où $\tilde{U} = \sum_{k=1}^{n_0} (-1)^{k-1} C_{n_0}^k \lambda^k U^k$ est manifestement un opérateur compact. Donc, d'après le théorème XII.1.4 déjà mentionné, les solutions de l'équation homogène

$$T_\lambda^{n_0}(x) = 0$$

forment un sous-espace de dimension finie, et $\mathrm{N}(T_\lambda) \subset \mathrm{N}(T_\lambda^{n_0})$.

b) Supposons par absurde qu'un disque $|\lambda| \leqslant r$ contient une infinité de valeurs caractéristiques. Extrayons de cet ensemble une suite $\{\lambda_n\}$ de valeurs caractéristiques distinctes: $|\lambda_n| \leqslant r$ $(n = 1, 2, \ldots)$. Soit $\{x_n\}$ la suite des éléments propres associés:

$$x_n - \lambda_n U(x_n) = 0 \quad (n = 1, 2, \ldots). \tag{4}$$

Prouvons (par récurrence) que les éléments $x_1, x_2, \ldots, x_n$ sont linéairement indépendants pour tout $n = 1, 2, \ldots$ Ceci est vrai pour $n = 1$. Supposons que cela est vrai pour un $n \geqslant 1$. Vérifions ceci pour les éléments $x_1, x_2, \ldots, x_n, x_{n+1}$. Supposons le contraire. On a alors

$$x_{n+1} = \sum_{k=1}^{n} \alpha_k x_k,$$

d'où en vertu de (4)

$$\frac{x_{n+1}}{\lambda_{n+1}} = \sum_{k=1}^{n} \frac{\alpha_k}{\lambda_k} x_k.$$

En combinant avec l'égalité précédente, on trouve

$$\sum_{k=1}^{n} \left(1 - \frac{\lambda_{n+1}}{\lambda_k}\right) \alpha_k x_k = 0.$$

Ceci exprime que les éléments $x_1, x_2, \ldots, x_n$ sont linéairement dépendants, puisque $1 - \frac{\lambda_{n+1}}{\lambda_k} \neq 0$ $(k = 1, 2, \ldots, n)$. Ce qui contredit l'hypothèse de récurrence.

Formons l'ensemble $\mathrm{X}_n = \mathscr{L}(\{x_1, x_2, \ldots, x_n\})$. Comme $\mathrm{X}_{n+1} \neq \mathrm{X}_n$ d'après ce qui a été démontré, le lemme de la quasi-perpendiculaire (IV.1.7, tome 1) affirme l'existence d'éléments $y_1, y_2, \ldots, y_n$, tels que

$$y_n \in \mathrm{X}_n, \quad \|y_n\| = 1, \quad \rho(y_{n+1}, \mathrm{X}_n) > 1/2 \quad (n = 1, 2, \ldots). \tag{5}$$

Si $x \in X_n$, c'est-à-dire $x = \sum_{k=1}^{n} \beta_k x_k$, alors $U(x) = \sum_{k=1}^{n} \frac{\beta_k}{\lambda_k} x_k \in X_n$.
Par ailleurs

$$T_{\lambda_n}(x) = x - \lambda_n U(x) = \sum_{k=1}^{n} \beta_k \left(1 - \frac{\lambda_n}{\lambda_k}\right) x_k \in X_{n-1}$$

(puisque le coefficient en $x_n$ est nul).

Supposons que $m > n$. Considérons l'expression

$$U(\lambda_m y_m) - U(\lambda_n y_n) = y_m - [T_{\lambda_m}(y_m) + U(\lambda_n y_n)] = y_m - \tilde{y}.$$

On a démontré que

$$T_{\lambda_m}(y_m) \in X_{m-1} \quad \text{et} \quad U(\lambda_n y_n) \in X_n \subset X_{m-1}.$$

Donc

$$\tilde{y} = T_{\lambda_m}(y_m) + U(\lambda_n y_n) \in X_{m-1}.$$

En vertu de (5)

$$\| U(\lambda_m y_m) - U(\lambda_n y_n) \| = \| y_m - \tilde{y} \| > 1/2,$$

ce qui toutefois contredit la compacité de l'opérateur $U$, puisque la suite $\{\lambda_n y_n\}$ est bornée ($\| \lambda_n y_n \| \leqslant r$).

c) On a $\lambda_1 U(x_1) = x_1$ et $\lambda_2 U^*(g_2) = g_2$. Donc,

$$g_2(x_1) = g_2(\lambda_1 U(x_1)) = \lambda_1 U^*(g_2)(x_1) = \lambda_1 \left(\frac{g_2}{\bar{\lambda}_2}\right)(x_1) = \frac{\lambda_1}{\bar{\lambda}_2} g_2(x_1),$$

ce qui n'est possible que si $g_2(x_1) = 0$, puisque $\lambda_1 \neq \bar{\lambda}_2$.

Notons enfin en conclusion que si $U$ est un opérateur compact dans un espace X de dimension infinie, le point 0 appartient au spectre de $U$.

## § 4. Résolvante

**4.1.** Nous poursuivons l'étude de l'équation

$$x - \lambda U(x) = y \tag{1}$$

en nous intéressant maintenant seulement au cas où elle admet une solution unique.

Soit $\lambda \neq 0$ une valeur non singulière de l'opérateur $U$. L'opérateur $B_\lambda$ défini à partir de la relation

$$I + \lambda B_\lambda = (I - \lambda U)^{-1} \tag{2}$$

s'appelle *résolvante* de l'opérateur $U$. Pour $\lambda = 0$ nous convenons que $B_0 = U$.

Si l'on considère le spectre et respectivement l'ensemble des valeurs régulières, au lieu de $B_\lambda$ il est plus commode d'envisager

l'opérateur

$$R_\mu = (\mu I - U)^{-1}, \tag{3}$$

qui a un sens pour toutes les valeurs régulières de l'opérateur $U$. L'opérateur $R_\mu$ sera appelé *résolvante* aussi. Aucune confusion de ces deux notions n'est à craindre, car il ressortira du contexte de laquelle d'entre elles il est question ; en outre en dernier recours nous aurons les indices pour les distinguer. Notons que la résolvante $B_\lambda$ se rencontre fréquemment en théorie des équations intégrales, où l'on l'appelle résolvante de Fredholm ; en analyse fonctionnelle la résolvante est $R_\mu$ (cf. 3.1).

Si $\mu \neq 0$, on a de toute évidence

$$R_\mu = \frac{1}{\mu} I + \frac{1}{\mu^2} B_{1/\mu}. \tag{4}$$

Inversement, de

$$(I + \lambda B_\lambda)(I - \lambda U) = (I - \lambda U)(I + \lambda B_\lambda) = I$$

on déduit

$$B_\lambda = U(I - \lambda U)^{-1} = (I - \lambda U)^{-1} U. \tag{5}$$

Donc, pour $\lambda \neq 0$

$$B_\lambda = \frac{1}{\lambda} U R_{1/\lambda} = \frac{1}{\lambda} R_{1/\lambda} U. \tag{6}$$

Les relations (4) et (6) permettent de généraliser à $R_\mu$ toutes les propositions prouvées pour $B_\lambda$ et inversement.

4.2. Etudions le comportement de la résolvante $B_\lambda$ pour de petits $\lambda$. Considérons la série

$$I + \lambda U + \lambda^2 U^2 + \ldots + \lambda^n U^n + \ldots \tag{7}$$

Si elle est convergente dans l'espace des opérateurs $B$ (**X**, **X**), d'après la remarque qui suit le théorème de Banach (V.4.5, tome 1), sa somme sera égale à $(I - \lambda U^{-1})$, c'est-à-dire

$$(I - \lambda U)^{-1} = I + \lambda U + \ldots + \lambda^n U^n + \ldots,$$

d'où, en vertu de (5),

$$B_\lambda = U + \lambda U^2 + \ldots + \lambda^n U^{n+1} + \ldots \tag{8}$$

Cette formule est valable pour les $\lambda$ tels que la série (7) soit convergente. Mais on a établi au V.4.2 (tome 1) que la série (7) est convergente si

$$\lim_{n\to\infty} \sqrt[n]{\|\lambda^n U^n\|} < 1$$

et divergente si

$$\lim_{n\to\infty} \sqrt[n]{\| \lambda^n U^n \|} > 1.$$

On est donc conduit au théorème suivant.

THÉORÈME 1. *La résolvante* $B_\lambda$ *se décompose en une série* (8) *suivant les puissances de* $\lambda$, *dont le rayon de convergence*

$$r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\| U^n \|}} = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\| B_0^n \|}}.$$

Si, grâce à la relation (4), on passe de la résolvante $B_\lambda$ à la résolvante $R_\mu$, on obtient le

COROLLAIRE. *La résolvante* $R_\mu$ *se décompose en série suivant les puissances de* $\mu^{-1}$:

$$R_\mu = \frac{1}{\mu} I + \frac{1}{\mu^2} U + \ldots + \frac{1}{\mu^n} U^{n-1} + \ldots \quad \left( |\mu| > \frac{1}{r} = \lim_{n\to\infty} \sqrt[n]{\| U^n \|} \right).$$

**4.3.** On peut indiquer une autre expression pour le rayon de convergence de la série (8), liée avec la disposition de l'ensemble caractéristique sur le plan complexe.

Commençons par prouver deux propositions auxiliaires.

LEMME 1. *Quels que soient* $\lambda$, $\mu \in \pi(U)$, *on a*

$$B_\lambda - B_\mu = (\lambda - \mu) B_\mu B_\lambda. \tag{9}$$

DÉMONSTRATION. De (5) il suit

$$B_\lambda - B_\mu = U (I - \lambda U)^{-1} - (I - \mu U)^{-1} U.$$

En multipliant à droite par $I - \lambda U$ et ensuite à gauche par $I - \mu U$, on obtient

$$(I - \mu U)(B_\lambda - B_\mu)(I - \lambda U) = (I - \mu U) U - U (I - \lambda U) = (\lambda - \mu) U^2,$$

donc

$$B_\lambda - B_\mu = (\lambda - \mu)(I - \mu U)^{-1} U \cdot U (I - \lambda U)^{-1} = (\lambda - \mu) B_\mu B_\lambda,$$

c.q.f.d.

COROLLAIRE. *Les opérateurs* $B_\lambda$ *et* $B_\mu$ *commutent, c'est-à-dire* $B_\lambda B_\mu = B_\mu B_\lambda$.

On démontre de façon analogue que pour tous les $\lambda$, $\mu \in \rho(U)$

$$R_\lambda - R_\mu = -(\lambda - \mu) R_\lambda R_\mu.$$

LEMME 2. *La résolvante $B_\lambda$ est une fonction continue du paramètre $\lambda$ en tout point de l'ensemble $\pi(U)$, c'est-à-dire si $\lambda_n \to \lambda_0$ ($\lambda_n$, $\lambda_0 \in \in \pi(U)$), alors $B_{\lambda_n} \to B_{\lambda_0}$.*

DÉMONSTRATION. Prouvons tout d'abord que la fonction réelle $\|B_\lambda\|$ est continue sur $\pi(U)$. Si $U = 0$, alors $B_\lambda = 0$ et notre proposition est prouvée. Si $U \neq 0$, alors $B_\lambda \neq 0$ et l'on peut prouver la continuité de la fonction $\frac{1}{\|B_\lambda\|}$. De (9) il suit

$$|\|B_\lambda\| - \|B_\mu\|| \leqslant \|B_\lambda - B_\mu\| = |\lambda - \mu| \, \|B_\mu B_\lambda\| \leqslant \leqslant |\lambda - \mu| \, \|B_\mu\| \, \|B_\lambda\|.$$

Donc

$$\left| \frac{1}{\|B_\mu\|} - \frac{1}{\|B_\lambda\|} \right| \leqslant |\lambda - \mu|,$$

ce qui montre que $\frac{1}{\|B_\lambda\|}$ est continue.

Prouvons maintenant la continuité de $B_\lambda$. L'ensemble $\pi(U)$ étant ouvert et $\lambda_0 \in \pi(U)$, il existe un disque $|\lambda - \lambda_0| \leqslant \varepsilon$ entièrement contenu dans $\pi(U)$. La fonction continue $\|B_\lambda\|$ est bornée dans ce disque; supposons par exemple que

$$\|B_\lambda\| \leqslant M \quad (|\lambda - \lambda_0| \leqslant \varepsilon). \tag{10}$$

D'après (9) et (10)

$$\|B_\lambda - B_{\lambda_0}\| \leqslant |\lambda - \lambda_0| \, \|B_{\lambda_0}\| \, \|B_\lambda\| \leqslant M^2 \, |\lambda - \lambda_0| \quad (|\lambda - \lambda_0| \leqslant \varepsilon),$$

ce qui prouve le lemme.

THÉORÈME 2. *Le rayon de convergence $r$ de la série (8) est égal à la distance $r_0$ du point $\lambda = 0$ à l'ensemble caractéristique $\chi(U)$.*

DÉMONSTRATION. La série (8) étant convergente dans le disque $|\lambda| < r$, donc la résolvante existant pour ces $\lambda$, le disque en question est contenu dans l'ensemble des valeurs non singulières. Donc $r \leqslant r_0$.

Prenons un élément $x \in \mathbf{X}$ et une fonctionnelle $f \in \mathbf{X}^*$ et considérons la fonction $\varphi$ de la variable complexe $\lambda$:

$$\varphi(\lambda) = f(B_\lambda(x)).$$

Prouvons que $\varphi$ est régulière sur l'ensemble $\pi(U)$. En effet, si $\lambda, \mu \in \in \pi(U)$, d'après (9) on obtient

$$\frac{\varphi(\mu) - \varphi(\lambda)}{\mu - \lambda} = \frac{f(B_\mu(x)) - f(B_\lambda(x))}{\mu - \lambda} = f(B_\mu B_\lambda(x)).$$

Le second membre admet la limite $f(B_\lambda^2(x))$ pour $\mu \to \lambda$ (lemme 2). Donc, il existe la dérivée continue

$$\varphi'(\lambda) = f(B_\lambda^2(x)).$$

Développons la fonction $\varphi$ en série de Taylor au voisinage de $\lambda_0 = 0$

$$\varphi(\lambda) = \varphi(0) + \frac{\varphi'(0)}{1!}\lambda + \ldots + \frac{\varphi^{(n)}(0)}{n!}\lambda^n + \ldots \qquad (11)$$

Ce développement a lieu dans un disque ne renfermant pas de points singuliers de la fonction $\varphi$, et *a fortiori* dans le disque $|\lambda| < r_0$. Mais d'après (8)

$$\varphi(\lambda) = \sum_{n=0}^{\infty} f(U^{n+1}(x))\lambda^n \quad (|\lambda| < r). \qquad (12)$$

Ceci étant, un théorème classique de la théorie des fonctions d'une variable complexe dit que les séries (11) et (12) sont identiques, de sorte que la série (12) est convergente pour $|\lambda| < r_0$.

Prenons un $\lambda_1 \in ]0, r_0[$. La convergence de la série (12) pour $\lambda = \lambda_1$ entraîne que

$$\lambda_1^n f(U^{n+1}(x)) \xrightarrow[n\to\infty]{} 0,$$

donc, puisque $f$ est arbitraire,

$$\lambda_1^n U^{n+1}(x) \to 0 \ (\sigma(X, X^*)).$$

Mais une suite faiblement convergente est bornée (VIII.1.1, tome 1):

$$\sup_n \|\lambda_1^n U^{n+1}(x)\| < \infty.$$

Cette inégalité étant réalisée pour chaque $x \in X$ et l'espace X étant complet, le théorème VII.1.1, tome 1, nous dit que

$$\sup \|\lambda_1^n U^{n+1}\| = M < \infty. \qquad (13)$$

Donc

$$\lambda_1 \lim_{n\to\infty} \sqrt[n]{\|U^n\|} \leqslant \lim_{n\to\infty} M^{1/n} = 1$$

et

$$\lambda_1 \leqslant \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\|U^n\|}} = r.$$

$r_0 \leqslant r$, puisque $\lambda_1$ peut être pris aussi proche que l'on veut de $r_0$. D'après l'inégalité prouvée plus haut, $r \leqslant r_0$, donc $r = r_0$, c.q.f.d.

Remarque 1. Soit $\lambda_0$ une valeur non singulière de l'opérateur $U$. Comme plus haut, on justifie le développement

$$B_\lambda = B_{\lambda_0} + (\lambda - \lambda_0) B_{\lambda_0}^2 + \ldots + (\lambda - \lambda_0)^n B_{\lambda_0}^{n+1} + \ldots,$$

qui a lieu dans le disque $|\lambda - \lambda_0| < \rho_0$, où $\rho_0$ est la distance du point $\lambda_0$ à l'ensemble caractéristique, ou, comme dans le théorème 1

$$\rho_0 = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\| B_{\lambda_0}^{n+1} \|}}.$$

En remplaçant la résolvante $B_\lambda$ par $R_\mu$, on obtient le résultat suivant.

COROLLAIRE 1. *Le développement*

$$R_\mu = \frac{1}{\mu} I + \frac{1}{\mu^2} U + \ldots + \frac{1}{\mu^{n+1}} U^n + \ldots \qquad (14)$$

*a lieu pour* $|\mu| > \frac{1}{r}$, *où* $1/r$ *est le rayon du plus petit disque centré en l'origine des coordonnées et contenant entièrement le spectre.*

Le nombre $1/r$ est appelé *rayon spectral* de l'opérateur $U$.

REMARQUE 2. Si $U$ est un opérateur auto-adjoint dans un espace hilbertien, alors $1/r = \| U \|$ d'après IX.5.3 (tome 1). En combinant ce résultat avec le corollaire du théorème 1, on arrive à l'intéressante relation

$$\| U \| = \lim_{n\to\infty} \sqrt[n]{\| U^n \|} \text{ *)}.$$

CORROLAIRE 2. *Le spectre $\sigma(U)$ d'un opérateur linéaire continu $U$ dans un B-espace complexe n'est pas vide.*

DEMONSTRATION. Si $\sigma(U) = \varnothing$, en tenant compte de la relation entre $\sigma(U)$ et $\chi(U)$ et du fait que $0 \notin \chi(U)$ on obtient que l'ensemble $\pi(U)$ des valeurs non singulières est le plan complexe tout entier. L'opérateur $B_\lambda \neq 0$ pour tout $\lambda$, puisqu'on peut admettre que $U \neq 0$. Supposons que $y_0 = B_{\lambda_0}(x_0) \neq 0$. Prenons $f \in X^*$, $f(y_0) \neq$ $\neq 0$. Comme dans la démonstration du théorème 2, on trouve que la fonction

$$\varphi(\lambda) = f(B_\lambda(x_0))$$

est régulière sur le plan complexe tout entier. Par ailleurs,

$$|\varphi(\lambda)| = |f(B_\lambda(x_0))| \leqslant \|f\| \, \|B_\lambda\| \, \|x_0\|.$$

Comme $\sigma(U) = \varnothing$, il existe $U^{-1}$ continu, d'où comme dans le lemme 2 on déduit que $\| R_{1/\lambda} \| \to \| U^{-1} \|$ pour $\lambda \to \infty$. Donc, en vertu de (6)

$$\| B_\lambda \| \leqslant 1/|\lambda| \, \| R_{1/\lambda} \| \, \| U \| \xrightarrow[\lambda\to\infty]{} 0.$$

---

*) Du reste, cette égalité découle de la relation plus forte

$$\| U \| = \sqrt[n]{\| U^n \|} \quad (n = 1, 2, \ldots),$$

que l'on peut obtenir en se servant du théorème du spectre de l'opérateur $\varphi(U)$ (théorème IX.5.4, tome 1) appliqué à la fonction $\varphi(t) = t^n$ ($n = 1, 2, \ldots$).

Par suite $\varphi(\lambda)$ est bornée et d'après le théorème de Liouville $\varphi(\lambda)$ est identiquement égale à une constante qui manifestement ne peut être que zéro. Or, $\varphi(\lambda_0) = f(y_0) \neq 0$, d'où contradiction.

On remarquera que le spectre de l'opérateur $U$ aurait pu être vide si on avait tenté de le définir dans un espace réel comme on l'a fait dans 3.1. C'est pour cette raison essentielle que nous considérons le cas complexe.

4.4. Appliquons le résultat obtenu à l'étude de la convergence de la méthode des approximations successives pour l'équation

$$x - U(x) = y. \tag{15}$$

On sait (cf. V.5.1, tome 1) que la convergence de la série

$$I + U + U^2 + \ldots + U^n + \ldots \tag{16}$$

assure la convergence de la méthode des approximations successives quelle que soit l'approximation initiale $x_0$. En faisant $\mu = 1$ dans (14), on arrive au critère suivant de convergence de la méthode d'approximations successives pour l'équation (15).

Théorème 3. *Si le spectre de l'opérateur $U$ est contenu dans le disque $|\mu| < 1$, la méthode d'approximations successives pour l'équation* (15) *converge quels que soient $y \in \mathbf{X}$ et l'approximation initiale $x_0 \in \mathbf{X}$. Si l'extérieur du disque $|\mu| \leqslant 1$ contient des points du spectre, il existe un ensemble $E \subset \mathbf{X}$, résidu* *) *dans* $\mathbf{X}$, *tel que le processus d'approximations successives pour l'équation* (15), *d'approximation initiale $x_0 = 0$, diverge pour $y \in E$.*

Demonstration. Il suffit de prouver simplement la deuxième partie du théorème. On remarquera que dans ce cas on a

$$\lim_{n\to\infty} \|U^n\| = \infty$$

et, par suite de la remarque qui suit le théorème VII.1.1, tome 1,

$$\sup_n \|U^n(y)\| = \infty \tag{17}$$

pour tous les $y \in \mathbf{X}$ à l'exception peut-être d'un ensemble $G$ de première catégorie de $\mathbf{X}$. Or, la convergence du processus d'approximations successives commencé à partir de $x_0 = 0$ équivaut à la convergence de la série

$$y + U(y) + U^2(y) + \ldots + U^n(y) + \ldots$$

et en vertu de (17) cette série est convergente si $y \notin G$.

Remarque 1. Si l'opérateur $U$ est tel que tous les points du spectre distincts de 0 sont des valeurs propres, on peut énoncer le théorème

*) On rappelle qu'un ensemble $E$ d'un espace $\mathbf{X}$ est un résidu si son complémentaire est de première catégorie dans $\mathbf{X}$ (cf. I.4.7, tome 1).

4—01292

sous une forme plus précise. Plus exactement, pour que la méthode d'approximations successives soit convergente, il faut et il suffit que toutes les valeurs propres de l'opérateur $U$ soient contenues dans le disque $|\mu| < 1$.

En effet, d'après le théorème prouvé, si la méthode d'approximations successives est convergente, le spectre de l'opérateur $U$ est contenu dans le disque $|\mu| \leqslant 1$. Si l'on admet qu'il existe une valeur propre $\mu_0$ située sur le cercle $|\mu| = 1$, alors, en faisant $y = y'$ dans l'équation (15), où $y'$ est l'élément propre associé à la valeur propre $\mu_0$, et en prenant $x_0 = 0$, on obtient pour la $n$-ième approximation $x_n$ l'expression

$$x_n = (1 + \mu_0 + \ldots + \mu_0^{n-1})\, y',$$

qui n'admet pas de limite.

Remarque 2. Les résultats du théorème et de la remarque 1 se simplifient considérablement si dans l'équation (15) l'opérateur $U$ est un opérateur auto-adjoint dans un espace hilbertien. Compte tenu de la remarque 2 de 4.3 on obtient dans ce cas: pour que la méthode d'approximations successives soit convergente il suffit que $\| U \| < 1$. Si tous les points du spectre distincts de 0 sont valeurs propres, alors cette condition est aussi nécessaire.

Formulons enfin le théorème 3 en termes d'ensemble caractéristique.

Théorème 3′. *Si l'ensemble caractéristique de l'opérateur $U$ est situé à l'extérieur du disque $|\lambda| \leqslant 1$, alors la méthode des approximations successives pour l'équation* (15) *est convergente quels que soient $y \in \mathrm{X}$ et l'approximation initiale $x_0 \in \mathrm{X}$. Si le disque $|\lambda| < 1$ renferme des points de l'ensemble caractéristique, il existe un ensemble $E \subset \mathrm{X}$, résidu dans* X, *tel que si $y \in E$, alors le processus d'approximations successives commencé à partir de $x_0 = 0$ est divergent.*

Ce théorème est justiciable de remarques analogues aux remarques 1 et 2. Nous les laissons au soin du lecteur.

4.5. Si $U$ est un opérateur compact, on peut ajouter aux résultats des théorèmes 1 et 2 des résultats plus fins concernant le comportement de la résolvante au voisinage d'une valeur caractéristique.

On dira qu'un opérateur linéaire continu $V$ est *de dimension finie* s'il applique un espace X dans un sous-espace de dimension finie $\tilde{\mathrm{X}} \subset \mathrm{X}$. Choisissons dans $\tilde{\mathrm{X}}$ un système complet d'éléments linéairement indépendants $x_1, x_2, \ldots, x_n$. Par définition, pour tout $x \in \mathrm{X}$

$$V(x) = \sum_{k=1}^{n} \alpha_k x_k.$$

Les coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ dépendent manifestement de $x$. En posant $\alpha_k = f_k(x)$ $(k = 1, 2, \ldots, n)$ on s'assure immédiatement que les fonctionnelles $f_k$ $(k = 1, 2, \ldots, n)$ sont linéaires et continues:

linéaires de toute évidence, continues, car si une suite d'éléments d'un espace normé de dimension finie tend vers 0, chaque coordonnée tend vers 0 (cf. IV.1.6, tome 1). On obtient donc

$$V(x) = f_1(x)\,x_1 + f_2(x)\,x_2 + \ldots + f_n(x)\,x_n. \tag{18}$$

Inversement, tout opérateur $V$ représentable sous la forme (18) sera visiblement de dimension finie.

On notera qu'un opérateur de dimension finie est nécessairement compact.

Considérons maintenant un opérateur compact $U$ et soit $\lambda_0$ une de ses valeurs caractéristiques. On a le

THÉORÈME 4. *L'opérateur $U$ admet la représentation*

$$U = U' + U'',$$

*où $U''$ est un opérateur de dimension finie, $U'$ un opérateur compact. Ceci étant, l'ensemble caractéristique de l'opérateur $U''$ est composé du seul point $\lambda_0$, quant à l'ensemble caractéristique de l'opérateur $U'$, il se déduit de celui de $U$ par élimination du point $\lambda_0$ de ce dernier.*

DÉMONSTRATION. On peut admettre que $\lambda_0 = 1$ (sinon on considère l'opérateur $\frac{1}{\lambda_0} U$ au lieu de $U$). Prouvons que pour $\lambda_0 = 1$, la représentation de l'opérateur $U$ par la somme $U' + U''$ indiquée dans le théorème 1.1 vérifie les hypothèses du théorème envisagé.

Nous utilisons les notations du théorème 1.1. Vérifions que l'opérateur $U''$ possède une valeur caractéristique unique: $\lambda_0 = 1$. En effet, si $x_0 \in \mathrm{N}(T)$, *a fortiori* $x_0 \in \mathrm{N}(T^r) = \mathrm{X}''$, et, par suite, par définition de l'opérateur $U''$

$$U''(x_0) = U(x_0) = x_0$$

et $\lambda_0 = 1$ est valeur caractéristique de l'opérateur $U''$.

Si

$$\lambda U''(x) = x \tag{19}$$

pour un $x \in \mathrm{X}$ ($x \neq 0$), alors $x \in \mathrm{X}''$ puisque $U''(x) \in \mathrm{X}''$ et, par suite, $U''(x) = U(x)$. Supposons que $m \geqslant 0$ est tel que $T^{m+1}(x) = 0$ mais $T^m(x) \neq 0$. En vertu de (19)

$$0 = T^m(x - \lambda U(x)) = T^m(x - \lambda[x - T(x)]) =$$
$$= (1 - \lambda)\,T^m(x) + \lambda T^{m+1}(x) = (1 - \lambda)\,T^m(x),$$

ce qui n'est possible que pour $\lambda = 1$.

Donc, l'unique valeur caractéristique de l'opérateur $U''$ est $\lambda_0 = 1$.

Prouvons maintenant la partie du théorème relative à l'ensemble caractéristique de l'opérateur $U'$.

L'opérateur $T' = I - U'$ admettant un inverse d'après le théorème 1.1, la valeur $\lambda_0 = 1$ n'est pas valeur caractéristique de $U'$. Soient $\lambda \neq 1$ une valeur caractéristique de l'opérateur $U$ et $x_0 \neq 0$ l'élément propre associé. Si $x_0 \in X''$, en raisonnant comme plus haut on obtiendrait que $\lambda = 1$. Donc $x_0 \notin X''$. Par suite, dans la décomposition

$$x_0 = x_0' + x_0'' \quad (x_0' \in X', \quad x_0'' \in X'') \tag{20}$$

(cf. c) théorème 1.1) on doit avoir $x_0' \neq 0$. En vertu de l'unicité de la décomposition (20) il vient de la relation

$$x_0 = \lambda U(x_0') + \lambda U(x_0'')$$

que

$$x_0' = \lambda U(x_0') = \lambda U'(x_0'),$$

c'est-à-dire $\lambda$ est valeur caractéristique de l'opérateur $U'$.

Inversement, supposons que $\lambda$ est valeur caractéristique de l'opérateur $U'$ et $x \neq 0$ l'élément propre correspondant. Comme

$$x = \lambda U'(x) \in X',$$

alors $U'(x) = U(x)$ et par suite $x = \lambda U(x)$, c'est-à-dire $\lambda$ est valeur caractéristique de $U$.

Les autres propositions du théorème sont contenues dans le théorème 1.1.

Le théorème est prouvé.

4.6. Le théorème suivant donne une idée plus complète sur le comportement de la résolvante au voisinage d'une valeur caractéristique.

THÉORÈME 5. *Soit $\lambda_0$ une valeur caractéristique d'un opérateur compact $U$. Alors, dans un voisinage assez petit du point $\lambda_0$, on a le développement*

$$B_\lambda = \frac{U_{-r}}{(\lambda-\lambda_0)^r} + \ldots + \frac{U_{-1}}{\lambda-\lambda_0} + U_0 + \\ + U_1(\lambda-\lambda_0) + \ldots + U_n(\lambda-\lambda_0)^n + \ldots, \tag{21}$$

*$r$ est le rang de la valeur caractéristique $\lambda_0$; les opérateurs $U_{-r}, \ldots, U_{-1}$ sont de dimension finie; l'opérateur $U_{-r} \neq 0$.*

*La série* (21) *converge dans l'espace des opérateurs $B$* (X, X).

DÉMONSTRATION. Comme dans la démonstration du théorème précédent, on admettra que $\lambda_0 = 1$. On remarquera aussitôt qu'en vertu du lemme 2 de 1.1, le rang de la valeur caractéristique $\lambda_0$ est fini. En utilisant encore les notations du théorème 1.1 on aura, en vertu de la remarque suivant ce dernier, $X' = T^r(X)$, $X'' = N(T^r)$.

Mettons l'élément $x \in X$ sous la forme

$$x = x' + x'' \quad (x' \in X', \; x'' \in X'')$$

(cf. c) du théorème 1.1), associons à $x$ l'élément $x' = P'(x)$ ou l'élément $x'' = P''(x)$ et construisons les opérateurs $P'$ et $P''$, projecteurs de l'espace X sur les sous-espaces X' et X''. Ces opérateurs sont continus en vertu de la majoration (9) du § 1. Signalons encore que

$$P'(\mathrm{X}) = \mathrm{X}', \quad P''(\mathrm{X}) = \mathrm{X}'', \quad P' + P'' = I.$$

Soit $y \in \mathrm{X}$. L'élément $x = B_\lambda(y)$ est solution de l'équation

$$x - \lambda U(x) = U(y). \tag{22}$$

En substituant dans cette équation $x = P'(x) + P''(x)$, $y = P'(y) + P''(y)$ et en tenant compte de ce que $U(\mathrm{X}') \subset \mathrm{X}'$ et $U(\mathrm{X}'') \subset \mathrm{X}''$, on peut mettre l'équation (22) sous la forme d'un système de deux équations:

$$\left.\begin{aligned} P'(x) - \lambda UP'(x) = UP'(y) \\ P''(x) - \lambda UP''(x) = UP''(y) \end{aligned}\right\}. \tag{23}$$

En remarquant que $UP' = U'P'$, on peut mettre la première de ces équations sous la forme

$$T'_\lambda(P'(x)) = U'P'(y),$$

où $T'_\lambda = I - \lambda U'$. Le théorème 4 nous dit que $\lambda_0 = 1$ est valeur régulière de l'opérateur $U'$. Donc, en vertu de la remarque suivant le théorème 2, si la différence $\lambda - 1$ est assez petite, la résolvante $B'_\lambda$ de l'opérateur $U'$ admet la décomposition

$$B'_\lambda = U'_0 + (\lambda - 1) U'_1 + \ldots + (\lambda - 1)^n U'_n + \ldots,$$

de plus, la série de droite converge dans l'espace $B$ (X, X). D'après ce qui vient d'être dit on peut écrire

$$P'(x) = B'_\lambda P'(y) =$$
$$= [U_0 + (\lambda - 1) U_1 + \ldots + (\lambda - 1)^n U_n + \ldots](y), \tag{24}$$

où $U_n = U'_n P'$ $(n = 1, 2, \ldots)$, et la série de droite converge de nouveau dans l'espace $B$ (X, X).

Voyons maintenant la deuxième équation (23).

Formons l'espace quotient

$$\mathrm{X}^{(k)} = \mathrm{N}(T^k)/\mathrm{N}(T^{k-1}) \quad (k = 1, 2, \ldots, r)$$

et désignons par $\varphi_k$ l'homomorphisme naturel de l'espace $\mathrm{N}(T^k)$ sur $\mathrm{X}^{(k)}$. L'espace $\mathrm{X}^{(r)}$ est visiblement de dimension finie. Choisissons dans cet espace un système complet d'éléments linéairement indépendants $\bar{x}_1^{(r)}, \bar{x}_2^{(r)}, \ldots, \bar{x}_{n_r}^{(r)}$ et soient $x_1^{(r)}, x_2^{(r)}, \ldots, x_{n_r}^{(r)}$ des éléments de $\mathrm{N}(T^r)$, tels que $\varphi_r(x_j^{(r)}) = \bar{x}_j^{(r)}$. Les éléments $x_j^{(r-1)} = T(x_j^{(r)})$ $(j = 1, 2, \ldots, n_r)$ appartiennent à $\mathrm{N}(T^{r-1})$. De

plus leurs images $\bar{x}_j^{(r-1)} = \varphi_{r-1}(x_j^{(r-1)})$ sont linéairement indépendantes, car si $\sum_{j=1}^{n_r} \alpha_j \bar{x}_j^{(r-1)} = 0$, alors

$$\sum_{j=1}^{n_r} \alpha_j x_j^{(r-1)} = T\left(\sum_{j=1}^{n_r} \alpha_j x_j^{(r)}\right) \in \mathrm{N}(T^{r-2});$$

autrement dit $\sum_{j=1}^{n_r} \alpha_j x_j^{(r)} \in \mathrm{N}(T^{r-1})$, donc

$$\sum_{j=1}^{n_r} \alpha_j \bar{x}_j^{(r)} = \varphi_r\left(\sum_{j=1}^{n_r} \alpha_j x_j^{(r)}\right) = 0,$$

ce qui n'est possible que pour $\alpha_1 = \alpha_2 = \ldots = \alpha_{n_r} = 0$.

Ajoutons aux éléments $\bar{x}_1^{(r-1)}, \bar{x}_2^{(r-1)}, \ldots, \bar{x}_{n_r}^{(r-1)}$ les éléments $\bar{x}_{n_r+1}^{(r-1)}, \ldots, \bar{x}_{n_{r-1}}^{(r-1)} \in \mathrm{X}^{(r-1)}$ pour former une base dans $\mathrm{X}^{(r-1)}$. Choisissons ensuite $x_{n_r+1}^{(r-1)}, \ldots, x_{n_{r-1}}^{(r-1)} \in \mathrm{N}(T^{r-1})$ de telle sorte que $\bar{x}_j^{(r-1)} = \varphi_{r-1}(x_j^{(r-1)})$.

En poursuivant ce raisonnement on construit pour chaque $k = 1, 2, \ldots, r$ des éléments $x_1^{(k)}, x_2^{(k)}, \ldots, x_{n_k}^{(k)}$ tels que

$$x_j^{(k)} \in \mathrm{N}(T^k), \quad T(x_j^{(k)}) = x_j^{(k-1)}$$
$$(j = 1, 2, \ldots, n_k; \quad k = 1, 2, \ldots, r; \quad x_j^{(0)} = 0). \tag{25}$$

De plus, les éléments

$$\bar{x}_1^{(k)} = \varphi_k(x_1^{(k)}), \quad \bar{x}_2^{(k)} = \varphi_k(x_2^{(k)}), \ldots, \bar{x}_{n_k}^{(k)} = \varphi_k(x_{n_k}^{(k)})$$

forment pour chaque $k = 1, 2, \ldots, r$ une base dans $\mathrm{X}^{(k)}$.

Posons

$$\mathrm{X}_k = \mathscr{L}(\{x_1^{(k)}, x_2^{(k)}, \ldots, x_{n_k}^{(k)}\}) \quad (k = 1, 2, \ldots, r).$$

D'après (25)

$$T(\mathrm{X}_k) \subset \mathrm{X}_{k-1} \quad (k = 1, 2, \ldots, r; \quad \mathrm{X}_0 = \{0\}). \tag{26}$$

Prouvons que les éléments $\{x_j^{(k)}\}$ ($j = 1, 2, \ldots, n_k$; $k = 1, 2, \ldots, r$) forment dans $\mathrm{X}''$ un système complet d'éléments linéairement indépendants.

Soit

$$\sum_{k=1}^{r} \sum_{j=1}^{n_k} \lambda_j^{(k)} x_j^{(k)} = 0.$$

Comme $x_j^{(k)} \in N(T^{r-1})$ pour $k<r$, l'application de l'opérateur $\varphi_r$ donne

$$\sum_{k=1}^{r} \sum_{j=1}^{n_k} \lambda_j^{(k)} \varphi_r(x_j^{(k)}) = \sum_{j=1}^{n_r} \lambda_j^{(r)} \overline{x}_j^{(r)} = 0,$$

donc $\lambda_j^{(r)} = 0$ $(j=1, 2, \ldots, n_r)$. On établit de façon analogue que les autres coefficients $\lambda_j^{(k)}$ sont nus. Considérons maintenant un élément $x'' \in X''$ quelconque. L'élément $\varphi_r(x'') \in X^{(r)}$, donc il existe des coefficients $\alpha_1^{(r)}, \alpha_2^{(r)}, \ldots, \alpha_{n_r}^{(r)}$ tels que

$$\varphi_r(x'') = \sum_{j=1}^{n_r} \alpha_j^{(r)} \overline{x}_j^{(r)} = \varphi_r\left(\sum_{j=1}^{n_r} \alpha_j^{(r)} x_j^{(r)}\right).$$

Donc

$$x'' - \sum_{j=1}^{n_r} \alpha_j^{(r)} x_j^{(r)} \in N(T^{r-1}).$$

En poursuivant ce raisonnement, on obtient finalement pour des $\alpha_j^{(k)}$

$$x'' - \sum_{k=1}^{r} \sum_{j=1}^{n_r} \alpha_j^{(k)} x_j^{(k)} \in N(T^0) = \{0\}$$

et par suite $x'' = \sum_{k=1}^{r} \sum_{j=1}^{n_k} \alpha_j^{(k)} x_j^{(k)}$.

Soit $x$ un élément quelconque de X. L'élément $P''(x) \in X''$, donc

$$P''(x) = \sum_{k=1}^{r} \sum_{j=1}^{n_k} \alpha_j^{(k)}(x)\, x_j^{(k)},$$

les coefficients $\alpha_j^{(k)}$, comme indiqué dans 4.5, sont des fonctionnelles linéaires. Si l'on désigne

$$P_k(x) = \sum_{j=1}^{n_k} \alpha_j^{(k)}(x)\, x_j^{(k)} \quad (k=1, 2, \ldots, r),$$

d'après ce qui a été dit, $P_k$ sera un opérateur continu de X dans $X_k$, et de plus

$$P'' = P_1 + P_2 + \ldots + P_r, \tag{27}$$

$$P_m P_k = \begin{cases} 0, & m \neq k, \\ P_m, & m = k \end{cases} \quad (k, m = 1, 2, \ldots, r). \tag{28}$$

En remarquant que $UP'' = P''U$, mettons le second membre de la deuxième équation (23) sous la forme $P''U(y)$ et remplaçons l'opéra-

teur $P''$ par la somme $P_1 + P_2 + \ldots + P_r$:

$$\sum_{k=1}^{r} P_k(x) - \lambda \sum_{k=1}^{r} UP_k(x) = \sum_{k=1}^{r} P_k U(y).$$

En appliquant l'opérateur $P_m$ aux deux membres de cette égalité, on obtient, compte tenu de (28),

$$P_m(x) - \lambda \sum_{k=1}^{r} P_m U P_k(x) = P_m U(y) \quad (m = 1, 2, \ldots, r). \tag{29}$$

Or

$$UP_k(x) = P_k(x) - TP_k(x),$$

donc, en vertu de (26),

$$P_m U P_k(x) = \begin{cases} P_m(x), & k = m, \\ -TP_{m+1}(x), & k = m+1, \\ 0, & k \neq m, \ m+1 \end{cases}$$

$$(k, m = 1, 2, \ldots, r; \ P_{r+1} = 0).$$

Grâce à cette relation on peut mettre l'équation (29) sous la forme plus simple

$$P_r(x) - \lambda P_r(x) = P_r U(y), \tag{30}$$

$$P_m(x) - \lambda [P_m(x) - TP_{m+1}(x)] = P_m U(y)$$
$$(m = 1, 2, \ldots, r-1). \tag{31}$$

De (30) il résulte que

$$P_r(x) = \frac{P_r U(y)}{1-\lambda},$$

donc, en vertu de (31),

$$P_{r-1}(x) = \frac{\lambda}{1-\lambda} TP_r(x) + \frac{1}{1-\lambda} P_{r-1}U(y) =$$
$$= \frac{\lambda}{(1-\lambda)^2} TP_r U(y) + \frac{1}{1-\lambda} P_{r-1} U(y)$$

et, d'une façon générale, pour tout $m = 1, 2, \ldots, r$

$$P_m(x) = \sum_{k=0}^{r-m} \frac{\lambda^k}{(1-\lambda)^{k+1}} T^k P_{m+k} U(y).$$

En ajoutant les égalités obtenues, on aura

$$P''(x) = \sum_{m=1}^{r} P_m(x) = \sum_{m=1}^{r} \sum_{k=0}^{r-m} \frac{\lambda^k}{(1-\lambda)^{k+1}} T^k P_{m+k} U(y) =$$
$$= \sum_{k=0}^{r-1} \frac{\lambda^k}{(1-\lambda)^{k+1}} \sum_{m=1}^{r-k} T^k P_{m+k} U(y). \tag{32}$$

Décomposons $\frac{\lambda^k}{(1-\lambda)^{k+1}}$ en fractions élémentaires:

$$\frac{\lambda^k}{(1-\lambda)^{k+1}} = \sum_{j=0}^{k} \frac{c_j^{(k)}}{(\lambda-1)^{j+1}},$$

$c_j^{(k)}$ sont des constantes telles que $c_k^{(k)} = (-1)^{k+1}$. En portant ceci dans (32) on trouve

$$P''(x) = \sum_{s=1}^{r} \frac{U_{-s}(y)}{(\lambda-1)^s}, \tag{33}$$

où les opérateurs $U_{-s}$ $(s = 1, 2, \ldots, r)$ sont des combinaisons linéaires d'opérateurs de la forme $T^k P_{m+k} U$. Comme $P_{m+k}(\mathrm{X}) = \mathrm{X}_{m+k} \subset \mathrm{X}''$ et que $T(\mathrm{X}'') \subset \mathrm{X}''$ en vertu de b) du théorème 1.1 il vient que les opérateurs $U_{-s}$ appliquent X dans X″ et, par suite, sont de dimension finie. D'autre part, de l'égalité (32) on voit que

$$U_{-r} = (-1)^r T^{r-1} P_r U.$$

Donc, si par exemple $y = x_1^{(r)}$, alors en vertu de (25)

$$U_{-r}(y) = (-1)^r T^{r-1} P_r U(x_1^{(r)}) = (-1)^r T^{r-1} P_r (x_1^{(r)} - x_1^{(r-1)}) =$$
$$= (-1)^r T^{r-1}(x_1^{(r)}) = (-1)^r x_1^{(1)} \neq 0,$$

de sorte que $U_{-r} \neq 0$.

En additionnant les relations (24) et (33) on obtient la décomposition souhaitée de la résolvante $B_\lambda$.

Ceci achève la démonstration du théorème.

Remarque. Si $U$ est un opérateur auto-adjoint dans un espace hilbertien, le théorème peut être précisé, car dans ce cas $r = 1$ (cf. 3.1) et, par suite, dans la décomposition (21) on n'aura qu'un terme de puissance négative en $\lambda - \lambda_0$, savoir $(\lambda - \lambda_0)^{-1} U_{-1}$. Nous ne formulons pas en détail le résultat correspondant, car il l'est déjà (dans une forme plus forte) dans IX.4.5, tome 1.

## § 5. Alternative de Fredholm

Dans ce paragraphe nous indiquons les conditions que doit remplir un opérateur linéaire continu $T$ d'un $B$-espace X dans lui-même, pour être justiciable de l'alternative de Fredholm (cf. 1.4). Il se trouve en particulier que si une puissance $U^m$ d'un opérateur linéaire $U$ est un opérateur compact, l'alternative de Fredholm est valable pour $T = I - U$. Les résultats exposés ci-après sont dus à S. Nikolski [1].

5.1. Soient l'équation

$$T(x) = y \tag{1}$$

et son adjointe

$$T^* (g) = f. \tag{2}$$

Soient par ailleurs les équations homogènes correspondantes

$$T (x) = 0, \tag{3}$$

$$T^* (g) = 0. \tag{4}$$

On rappelle que dire que l'alternative de Fredholm est valable pour un opérateur $T$ revient à dire que

1) ou bien les équations (1) et (2) admettent une solution quels que soient les seconds membres et cette solution est unique;

2) ou bien les équations (3) et (4) possèdent le même nombre fini de solutions linéairement indépendantes $x_1, x_2, \ldots, x_n$ et $g_1, g_2, \ldots, g_n$ respectivement; dans ce cas, pour que l'équation (1), respectivement l'équation (2), admette une solution il est nécessaire et suffisant que

$$g_k (y) = 0 \quad (k = 1, 2, \ldots, n),$$

respectivement

$$f (x_k) = 0 \quad (k = 1, 2, \ldots, n);$$

ceci étant la solution générale de l'équation (1) est

$$x = x^* + \sum_{k=1}^{n} c_k x_k,$$

celle de l'équation (2)

$$g = g^* + \sum_{k=1}^{n} c_k g_k,$$

où $x^*$ (resp. $g^*$) est une solution de l'équation (1) (resp. (2)) et $c_1, c_2, \ldots, c_n$ des constantes arbitraires.

Le théorème suivant exprime que la classe des opérateurs $T$ justiciables de l'alternative de Fredholm diffère pratiquement peu de la classe des opérateurs de la forme $T = I - U$, où $U$ est un opérateur compact.

THÉORÈME 1. *Chacune des deux conditions suivantes est une condition nécessaire et suffisante pour que l'alternative de Fredholm ait lieu pour un opérateur $T$.*

1) *L'opérateur $T$ admet la décomposition*

$$T = W + V,$$

*où $W$ est un opérateur admettant un inverse continu, $V$ un opérateur compact.*

2) *L'opérateur $T$ admet la décomposition*

$$T = W_1 + V_1,$$

*où $W_1$ est un opérateur admettant un inverse continu, $V_1$ un opérateur de dimension finie.*

Démonstration. On peut de toute évidence se contenter de prouver la condition suffisante de 1) et la condition nécessaire de 2).

*Condition suffisante* de 1). Supposons que

$$T = W + V,$$

$W$ admettant un inverse continu $W^{-1}$, $V$ étant compact. L'équation (1) est équivalente alors à l'équation

$$W^{-1}T(x) = W^{-1}(y). \tag{5}$$

Il existe par ailleurs un opérateur inverse continu $W^{*^{-1}} = (W^{-1})^*$ (XII.2.3), donc l'équation (2) équivaut à l'équation

$$T^* W^{*^{-1}}(g) = f \tag{6}$$

au sens que si $g_0$ est solution de l'équation (6), $W^{*-1}(g_0)$ sera solution de l'équation (2) et si $g_0'$ est solution de (2), $W^*(g_0')$ sera solution de (6).

Posons $U = -W^{-1}V$. Comme $U^* = -V^*(W^{-1})^* = -V^*W^{*^{-1}}$ (cf. IX.3.1, tome 1) on peut mettre les équations (5) et (6) sous la forme

$$x - U(x) = W^{-1}(y), \tag{7}$$

$$g - U^*(g) = f. \tag{8}$$

L'opérateur $U$ étant compact, les conclusions du théorème 1.4 sont valables pour les équations (7) et (8). Donc les équations homogènes

$$x - U(x) = 0, \tag{9}$$

$$g - U^*(g) = 0 \tag{10}$$

possèdent le même nombre fini de solutions linéairement indépendantes $x_1, x_2, \ldots, x_n$ et $g_1', g_2', \ldots, g_n'$. L'équation homogène (3) admettra visiblement le même système complet de solutions linéairement indépendantes que l'équation (9), c'est-à-dire $x_1, x_2, \ldots, x_n$. Prouvons que les fonctionnelles

$$g_k = W^{*^{-1}}(g_k') \quad (k = 1, 2, \ldots, n) \tag{11}$$

forment un système complet de solutions linéairement indépendantes de l'équation (4). Que chaque fonctionnelle (11) soit solution de l'équation (4) découle de l'équivalence, mentionnée plus haut, des équations (2) et (6). Les fonctionnelles (11) sont linéairement indépendantes, car la relation $\sum_{k=1}^{n} \alpha_k g_k = 0$ entraîne que $\sum_{k=1}^{n} \alpha_k g_k' = \sum_{k=1}^{n} \alpha_k W^*(g_k) = 0$, or cela n'est possible que pour $\alpha_1 = \alpha_2 =$

$= \ldots = \alpha_n = 0$. Enfin, si l'équation (4) avait une solution $g_0$ qui ne soit pas combinaison linéaire des fonctionnelles (11), la fonctionnelle $W^*(g_0)$ serait solution de l'équation (10) et cette solution ne serait pas combinaison linéaire des fonctionnelles $g_1'$, $g_2'$, . . ., $g_n'$, ce qui est impossible.

Donc les équations (3) et (4) possèdent le même nombre fini de solutions linéairement indépendantes.

Le théorème 1.4 nous dit que l'équation (5), et donc l'équation (1), admet une solution si et seulement si

$$g_k'(W^{-1}(y)) = 0 \quad (k = 1, 2, \ldots, n). \tag{12}$$

D'après (11), cette condition équivaut à:

$$(W^{-1})^*(g_k')(y) = W^{*^{-1}}(g_k')(y) = g_k(y) = 0 \quad (k = 1, 2, \ldots, n).$$

On vérifie de façon analogue que les conditions

$$f(x_k) = 0 \quad (k = 1, 2, \ldots, n)$$

sont nécessaires et suffisantes pour que l'équation (2) admette une solution.

*Condition nécessaire* de 2). Soient $x_1, x_2, \ldots, x_n$ et $g_1, g_2, \ldots$ . . ., $g_n$ des systèmes complets de solutions linéairement indépendantes des équations (3) et (4) respectivement. En utilisant le théorème V.7.4, tome 1, et le lemme III.3.1, tome 1, on trouve des fonctionnelles $f_1, f_2, \ldots, f_n \in \mathbf{X}^*$ et des éléments $y_1, y_2, \ldots, y_n \in$ $\in \mathbf{X}$ tels que

$$f_j(x_k) = \begin{cases} 0, & j \neq k, \\ 1, & j = k \end{cases} \quad (j, k = 1, 2, \ldots, n), \tag{13}$$

$$g_k(y_j) = \begin{cases} 0, & j \neq k, \\ 1, & j = k \end{cases} \quad (j, k = 1, 2, \ldots, n). \tag{14}$$

Appelons $\mathbf{Y}' = T(\mathbf{X})$, $\mathbf{Y}'' = \mathscr{L}(\{y_1, y_2, \ldots, y_n\})$. Chaque élément $y \in \mathbf{X}$ peut être écrit de façon unique sous la forme

$$y = y' + y'' \quad (y' \in \mathbf{Y}', \quad y'' \in \mathbf{Y}''). \tag{15}$$

En effet, si l'on pose

$$y'' = \sum_{k=1}^{n} g_k(y)\, y_k, \quad y' = y - y'',$$

alors, d'après (14),

$$g_j(y') = g_j(y) - \sum_{k=1}^{n} g_k(y)\, g_j(y_k) = 0 \quad (j = 1, 2, \ldots, n),$$

si bien que l'équation $T(x) = y'$ admet une solution et, par suite, $y' \in \mathbf{Y}'$. L'unicité de la représentation (15) résulte du fait que si

$y'' = \sum_{k=1}^{n} \alpha_k y_k \in Y'$, alors l'équation $T(x) = y''$ admet une solution et, par suite, $g_j(y'') = \alpha_j = 0$ $(j = 1, 2, \ldots, n)$.

Appelons d'autre part $X' = N(f_1, f_2, \ldots, f_n)$ et $X'' = \mathscr{L}(\{x_1, x_2, \ldots, x_n\})$. Comme dans ce qui précède prouvons que tout élément $x \in X$ peut être représenté de manière unique sous la forme

$$x = x' + x'' \quad (x' \in X', \; x'' \in X''). \tag{16}$$

Construisons l'opérateur $W_1$ en admettant que

$$W_1(x) = T(x) + \sum_{k=1}^{n} f_k(x)\, y_k,$$

et montrons que $W_1$ réalise une bijection de X sur lui-même, donc possède un inverse continu (théorème XII.1.2). En effet, soit $y$ un élément quelconque de X. Mettons-le sous la forme (15):

$$y = y' + y'',$$

où $y' \in Y' = T(X)$, $y'' = \sum_{k=1}^{n} \alpha_k y_k \in Y''$, c'est-à-dire l'équation $T(x) = y'$ admet une solution $x'$ que l'on peut supposer appartenir à $X'$ *).

En posant

$$x'' = \sum_{k=1}^{n} \alpha_k x_k, \quad x^* = x' + x''$$

et en tenant compte de ce que $T(x'') = 0$, $f_j(x') = 0$ et de (13), on obtient

$$W_1(x^*) = T(x') + T(x'') + \sum_{k=1}^{n} f_k(x'')\, y_k = T(x') + \sum_{k=1}^{n} f_k\Big(\sum_{j=1}^{n} \alpha_j x_j\Big) y_k = y' + \sum_{k=1}^{n} \alpha_k y_k = y.$$

Montrons que l'équation $W_1(x) = y$ n'admet de solutions autres que $x^*$. En effet, dans le cas contraire il existerait un élément $x_0 \neq 0$ tel que

$$W_1(x_0) = 0,$$

c'est-à-dire

$$T(x_0) + \sum_{k=1}^{n} f_k(x_0)\, y_k = 0.$$

*) En effet, en mettant la solution $x$ sous la forme (16): $x = x' + x''$, et, en tenant compte de ce que $T(x'') = 0$, on aura $y' = T(x) = T(x')$.

De plus $T(x_0) \in \mathrm{Y}$ et $\sum_{k=1}^{n} f_k(x_0)\, y_k \in \mathrm{Y}''$. L'unicité de la représentation de l'élément $x_0$ sous la forme (15) nous conduit aux relations

$$T(x_0) = 0, \quad \sum_{k=1}^{n} f_k(x_0)\, y_k = 0, \quad f_k(x_0) = 0 \quad (k = 1, 2, \ldots, n),$$

d'où il résulte que $x_0 = 0$ (puisque $x_0$ appartient à la fois à $\mathrm{X}''$ et à $\mathrm{X}'$).

Pour achever la démonstration du théorème il suffit de poser

$$V_1(x) = -\sum_{k=1}^{n} f_k(x)\, y_k.$$

Remarque. On laisse au lecteur le soin de vérifier que si dans les conditions 1) ou 2), on remplace l'opérateur $T$ par $T^*$, on obtient deux conditions qui sont également nécessaires et suffisantes pour que l'opérateur $T$ soit justiciable de l'alternative de Fredholm.

5.2. La suite de l'exposé se base sur deux lemmes simples.

Lemme 1. *Soient $A$ et $B$ deux opérateurs linéaires continus d'un espace normé* X *dans lui-même. Si ces opérateurs commutent et si l'opérateur $C = AB$ est inversible, les opérateurs $A$ et $B$ le sont également.*

Démonstration. Prouvons tout d'abord que les opérateurs $A$ et $C^{-1}$ commutent. En effet, on a

$$A = C^{-1}CA = C^{-1}ABA = C^{-1}A(AB) = C^{-1}AC.$$

Une multiplication à droite par $C^{-1}$ nous donne $AC^{-1} = C^{-1}A$.

La permutabilité des opérateurs $A$ et $C^{-1}$ nous permet d'écrire

$$B(AC^{-1}) = BAC^{-1} = CC^{-1} = I$$

et

$$(AC^{-1})B = C^{-1}AB = C^{-1}C = I,$$

d'où résulte l'existence de $B^{-1} = AC^{-1}$. On démontre de façon analogue l'existence de $A^{-1} = BC^{-1}$.

Remarque. Si l'opérateur $C^{-1}$ est continu, il en sera de même des opérateurs $A^{-1}$ et $B^{-1}$.

Lemme 2. *Soit $U$ un opérateur continu dans un espace* X. *Entre l'ensemble caractéristique $\chi(U)$ de l'opérateur $U$ et l'ensemble caractéristique $\chi(U^m)$ de l'opérateur $U^m$ on a la relation*

$$[\chi(U)]^m \subset \chi(U^m),$$

*c'est-à-dire si $\lambda \in \chi(U)$, $\lambda^m \in \chi(U^m)$.*

Démonstration. Posons $\varepsilon = e^{2\pi i/m}$. On a

$$I - \lambda^m U^m = (I - \lambda U)(I - \lambda\varepsilon U) \ldots (I - \lambda\varepsilon^{m-1} U).$$

Si $\lambda^m \notin \chi(U^m)$, en posant

$$A = I - \lambda U, \quad B = (I - \lambda\varepsilon U) \ldots (I - \lambda\varepsilon^{m-1}U),$$
$$C = I - \lambda^m U^m,$$

on trouve qu'il existe l'inverse continu $C^{-1}$. Donc, d'après la remarque suivant le lemme 1, il existe l'inverse continu $A^{-1}$, c'est-à-dire $\lambda \notin \chi(U)$.

**5.3.** Admettons comme dans 5.1 que **X** est un $B$-espace et considérons un opérateur linéaire continu $U$ dans **X**.

THÉORÈME 2. *Supposons qu'il existe un entier naturel $m$ tel que l'opérateur $U^m$ soit compact. Alors l'alternative de Fredholm est valable pour l'opérateur $T = I - U$.*

DÉMONSTRATION. D'après le lemme 2, l'ensemble caractéristique $\chi(U)$ est composé de points isolés, donc le cercle unité du plan complexe ne renferme qu'un nombre fini de points $\lambda_1, \lambda_2, \ldots$ $\ldots, \lambda_\nu \in \chi(U)$.

Supposons que $p$ parcourt l'ensemble des nombres premiers; les nombres

$$e^{2k\pi i/p} \quad (k = 1, 2, \ldots, p-1)$$

sont tous distincts, donc pour $p_0$ assez grand

$$\lambda_j \neq e^{2k\pi i/p} \quad (p \geqslant p_0; \; k = 1, 2, \ldots, p-1; \; j = 1, 2, \ldots, \nu). \quad (17)$$

On peut admettre que $m$ est un nombre premier, et de plus $m \geqslant$ $\geqslant p_0$. Considérons le développement

$$I - U^m = (I - U)(I - \varepsilon U) \ldots (I - \varepsilon^{m-1}U) = (I - U)V, \quad (18)$$

où

$$\varepsilon = e^{2\pi i/m}, \quad V = (I - \varepsilon U) \ldots (I - \varepsilon^{m-1}U).$$

D'après (17), les opérateurs $I - \varepsilon^k U$ $(k = 1, 2, \ldots, m-1)$ sont inversibles, donc existe l'opérateur continu $V^{-1}$. Mais alors

$$T = I - U = (I - U^m)V^{-1} = V^{-1} - U^m V^{-1}.$$

L'opérateur $V^{-1}$ étant inversible et $U^m V^{-1}$ compact, on peut se servir du théorème 1.

Ceci achève la démonstration du théorème.

**5.4.** Le théorème de l'ensemble caractéristique d'un opérateur compact (théorème 2) est valable pour les opérateurs du théorème précédent. Plus exactement on a le

THÉORÈME 3. *Si l'opérateur $U^m$ est compact pour un $m$, alors*

*1) l'ensemble caractéristique $\chi(U)$ de l'opérateur $U$ est composé uniquement de valeurs caractéristiques, chaque valeur caractéristique étant de rang fini et le sous-espace propre associé de dimension finie;*

2) *chaque disque* $|\lambda| \leqslant R$ *du plan complexe ne contient qu'un nombre fini de valeurs caractéristiques.*

DEMONSTRATION. Grâce au lemme 2 on peut se contenter de prouver seulement la partie 1) du théorème. La première proposition de 1) découlant manifestement du théorème 2, il nous reste à démontrer que le rang de chaque valeur caractéristique est fini et que le sous-espace propre associé est de dimension finie.

Sans restreindre la généralité on peut admettre que la valeur caractéristique considérée est $\lambda_0 = 1$. Posons $T_m = I - U^m$. En vertu de (18) on a

$$T_m = VT, \quad T = V^{-1}T^m.$$

Les opérateurs $T$, $T_m$, $V$ et $V^{-1}$ commutant entre eux, on a

$$T_m^n = V^n T^n, \quad T^n = V^{-n} T_m^n \quad (n = 1, 2, \ldots).$$

D'où il suit immédiatement que

$$\mathbf{N}(T_m^n) = \mathbf{N}(T^n). \tag{19}$$

L'opérateur $U^m$ étant compact et la proposition à prouver ayant déjà été établie pour de tels opérateurs, en vertu de (19) elle est vraie pour le cas envisagé.

En conclusion citons un exemple d'opérateur linéaire continu $U$ non compact dont la puissance $U^2$ est compacte.

Soit $\mathbf{X}$ l'un des espaces $\mathbf{l}^p$ $(1 \leqslant p \leqslant \infty)$, $\mathbf{c}$, $\mathbf{c}_0$. Pour $x = \{\xi_n\} \in \mathbf{X}$ posons $y = U(x)$, où

$$y = \{\eta_n\} \left(\eta_n = \begin{cases} 0 & \text{pour } n \text{ impair,} \\ \xi_{n-1} & \text{pour } n \text{ pair;} \end{cases} \quad n = 1, 2, \ldots \right).$$

De toute évidence $U^2 = 0$.

REMARQUE. Les théorèmes du § 4, prouvés pour des opérateurs compacts, n'ayant utilisé que les propriétés de compacité des opérateurs mentionnées dans le théorème 1.1 et le théorème 3.1, et ces théorèmes s'étendant sans changement aux opérateurs examinés plus haut, les résultats du § 4 sont également valables si l'on considère qu'une puissance seulement de l'opérateur $U$ est compacte.

## § 6. Application aux équations intégrales

6.1. Considérons l'équation intégrale

$$x(s) - \lambda \int_0^1 K(s, t)\, x(t)\, dt = y(s), \tag{1}$$

où $K(s, t)$ est supposé continu dans le carré $[0, 1; 0, 1]$. Si l'on traite l'intégrale comme un opérateur linéaire dans l'espace $\mathbf{C}[0, 1]$,

on s'aperçoit que l'équation (1) est une équation du type étudié dans les paragraphes précédents.

On pourrait envisager une équation intégrale d'une forme plus générale que (1), par exemple

$$x(s)-\lambda \int_T K(s, t)\, x(t)\, dt = y(s), \tag{1'}$$

où $T$ est un ensemble borné fermé d'un espace euclidien à $n$ dimensions ($s$ et $t$ sont des points de cet espace). Toutes les démonstrations effectuées pour l'équation (1) étant valables sans changements notables pour l'équation (1'), nous étudierons le cas le plus simple.

L'opérateur intégral $U$

$$z = U(x), \quad z(s) = \int_0^1 K(s, t)\, x(t)\, dt, \tag{2}$$

traité comme un opérateur de $\mathbf{C}[0, 1]$ dans $\mathbf{C}[0, 1]$ a pour norme (cf. III.2.4, tome 1)

$$\|U\| = \max_s \int_0^1 |K(s, t)|\, dt$$

et est compact (IX.2.1, tome 1).

Ecrivons l'équation (1) sous la forme

$$x - \lambda U(x) = y. \tag{3}$$

La solution $x^*$ de cette équation, dont l'expression en fonction de $y$ est

$$x^* = y + \lambda B_\lambda(y),$$

peut être, d'après le théorème 4.1, développée en la série

$$x^* = y + \lambda U(y) + \ldots + \lambda^n U^n(y) + \ldots, \tag{4}$$

qui converge pour tous les

$$|\lambda| < \frac{1}{d} = r,$$

où $d = \lim\limits_{n\to\infty} \sqrt[n]{\|U^n\|}$, et $r$ la distance du point $\lambda = 0$ à l'ensemble caractéristique de l'opérateur $U$ (théorème 4.2). La série (4) est, en tout cas, convergente pour

$$|\lambda| < \frac{1}{\|U\|} = \frac{1}{\max\limits_s \int_0^1 |K(s, t)|\, dt}.$$

Comme indiqué dans V.3.8, tome 1, les puissances de l'opérateur $U$ seront aussi des opérateurs intégraux. Notamment:

$$z = U^n(x), \quad z(s) = \int_0^1 K_n(s, t)\, x(t)\, dt \quad (n = 1, 2, \ldots), \tag{5}$$

où $K_n(s, t)$ sont les noyaux itérés.

En portant (5) dans (4), on obtient le développement de la solution de l'équation intégrale (1) en série suivant les puissances du paramètre

$$x^*(s) = y(s) + \lambda \int_0^1 K(s, t)\, y(dt) + \ldots + \lambda^n \int_0^1 K_n(s, t)\, y(t)\, dt + \ldots$$

De plus, la série converge uniformément en $s \in [0, 1]$.

La série

$$B_\lambda = U + \lambda U^2 + \ldots + \lambda^{n-1} U^n + \ldots \tag{6}$$

étant convergente dans l'espace des opérateurs de $\mathbf{C}[0, 1]$ dans $\mathbf{C}[0, 1]$, on a

$$\left\| \sum_{j=m+1}^{m+p} \lambda^{j-1} U^j \right\| = \max_s \int_0^1 \left| \sum_{j=m+1}^{m+p} \lambda^{j-1} K_j(s, t) \right| dt \xrightarrow[m \to \infty]{} 0.$$

Donc, pour chaque $s \in [0, 1]$ fixe, la série

$$\sum_{j=1}^{\infty} \lambda^{j-1} K_j(s, t) \tag{7}$$

converge uniformément en $s \in [0, 1]$ dans l'espace $\mathbf{L}^1$. La somme de cette série, la fonction $\Gamma(s, t; \lambda)$, est appelée *résolvante* de l'équation intégrale (1). Il est clair que

$$B_\lambda(y)(s) = \int_0^1 \Gamma(s, t; \lambda)\, y(t)\, dt,$$

en conséquence de quoi on peut mettre (4) sous la forme

$$x^*(s) = y(s) + \lambda \int_0^1 \Gamma(s, t; \lambda)\, y(t)\, dt.$$

Si $|\lambda| < r$, le théorème 1.3 nous dit que le processus d'approximations successives pour l'équation (3) est convergent; appliqué à l'équation intégrale (1) ceci nous amène au résultat suivant: pour $|\lambda| < r$ la solution de l'équation (1) peut être obtenue comme la limite d'une suite $\{x_n(s)\}$ uniformément convergente de fonctions

continues, définies par la formule récurrentielle

$$x_{n+1}(s)=\lambda\int_0^1 K(s,\ t)\,x_n(t)\,dt+y(s)\quad (n=0,\ 1,\ \ldots),$$

$x_0(t)$ étant une fonction continue quelconque.

6.2. Si le noyau $K(s,\ t)$ s'annule pour $s<t$, l'équation (1) devient

$$x(s)-\lambda\int_0^s K(s,\ t)\,x(t)\,dt=y(s). \tag{8}$$

Les équations de cette nature sont appelées *équations intégrales de Volterra.*

Il est immédiat de vérifier que les noyaux itérés de l'équation de Volterra s'annulent également pour $s<t$.

Supposons que le noyau $K(s,\ t)$ est continu pour $0\leqslant t\leqslant$ $\leqslant s\leqslant 1$ et prouvons que le développement (4) a lieu pour tous les $\lambda$ complexes, c'est-à-dire $r=\infty$.

Posons $|K(s,\ t)|\leqslant M$. Pour $K_n(s,\ t)$ on a immédiatement

$$|K_n(s,\ t)|\leqslant\frac{s^{n-1}}{(n-1)!}M^n\quad (n=1,\ 2,\ \ldots). \tag{9}$$

En effet, cette majoration est triviale pour $n=1$, et si elle est valable pour un $n>1$, alors

$$|K_{n+1}(s,\ t)|\leqslant\int_0^s|K(s,\ u)\,K_n(u,\ t)|\,dt\leqslant$$

$$\leqslant M\int_0^s\frac{u^{n-1}}{(n-1)!}M^n\,du=\frac{s^n}{n!}M^{n+1}.$$

De (9) il résulte

$$\|U^n\|=\max_s\int_0^s|K_n(s,\ t)|\,dt\leqslant M^n\max_s\int_0^s\frac{s^{n-1}}{(n-1)!}\,dt=\frac{M^n}{(n-1)!},$$

d'où

$$\sqrt[n]{\|U^n\|}\leqslant M\sqrt[n]{\frac{1}{(n-1)!}}\xrightarrow[n\to\infty]{}0.$$

Donc, $r=\infty$ et l'opérateur intégral de type Volterra ne possède pas de valeurs caractéristiques.

6.3. Revenons à l'équation (1). L'opérateur (2) étant compact, l'alternative de Fredholm a lieu pour l'équation (3). Ceci nous conduit au résultat suivant pour l'équation intégrale (1).

THÉORÈME 1. *Ou bien l'équation* (1) *admet une solution unique continue quelle que soit la fonction continue* $y(s)$, *ou bien l'équation*

$$x(s)-\lambda\int_0^1 K(s,\ t)\,x(t)\,dt=0 \tag{10}$$

*possède un nombre fini de solutions linéairement indépendantes* $x_1(s)$, $x_2(s), \ldots, x_n(s)$. *Ceci étant, l'équation*

$$\psi(t)-\overline{\lambda}\int_0^1 \overline{K(s,\ t)}\,\psi(s)\,ds=0$$

*admet aussi n solutions continues linéairement indépendantes* $\psi_1(t)$, $\psi_2(t), \ldots, \psi_n(t)$. *Dans le dernier cas, l'équation* (1) *possède une solution si et seulement si*

$$\int_0^1 \overline{\psi_k(s)}\,y(s)\,ds=0.$$

Les valeurs de $\lambda$ pour lesquelles l'équation (10) possède des solutions non nulles sont appelées *valeurs caractéristiques* de l'équation (1) ou du noyau $K(s,\ t)$. Autrement dit, les valeurs caractéristiques de l'équation (1) ne sont autres que les valeurs caractéristiques de l'opérateur $U$. Donc, pour la plus petite valeur caractéristique en module de l'équation intégrale on a la minoration

$$|\lambda_1|=R=\frac{1}{d}\geqslant\frac{1}{\|U\|}=\frac{1}{\max\int_0^1|K(s,\ t)|\,dt}\geqslant\frac{1}{\max\limits_{s,\,t}|K(s,\ t)|}.$$

En utilisant la dépendance de la solution de l'équation (3) par rapport à $\lambda$, on arrive, en vertu du théorème 4.5, au résultat suivant.

THÉORÈME 2. *Au voisinage d'une valeur caractéristique* $\lambda_0$ *la solution de l'équation* (1) *peut être mise sous la forme*

$$x^*(s)=\frac{y_{-r}(s)}{(\lambda-\lambda_0)^r}+\ldots+\frac{y_{-1}(s)}{\lambda-\lambda_0}+ \\ +y_0(s)+\ldots+y_n(s)(\lambda-\lambda_0)^n+\ldots,$$

*où* $y_k(s)$ $(k=-r, \ldots, 0, \ldots)$ *sont des fonctions continues dépendant uniquement de* $y$. *La série du second membre est uniformément convergente en* $s\in[0,\ 1]$.

Donc, pour chaque $s\in[0,\ 1]$ fixe, $x^*(s)$ est une fonction méromorphe de $\lambda$, de pôles aux valeurs caractéristiques.

6.4. Considérons maintenant l'espace $L^p(D)$, où $D$ est un domaine borné d'un espace à $n$ dimensions, et l'équation intégrale

$$x(s)-\lambda\int_D K(s,\ t)\,x(t)\,dt=y(s)\quad (s\in D). \tag{11}$$

Supposons que le noyau $K(s, t)$ vérifie les conditions du théorème XI.3.3, tome 1 (pour $q=p$), c'est-à-dire

1) $\left\{\int_D |K(s,\ t)|^r\,dt\right\}^{1/r}\leqslant C_1\quad (s\in D);$

2) $\left\{\int_D |K(s,\ t)|^\sigma\,ds\right\}\leqslant C_2\quad (t\in D);$

3) $p-r(p-1)<\sigma<p.$

Le théorème cité ci-dessus nous dit que l'opérateur intégral $U$ de noyau $K(s, t)$ sera un opérateur compact de $L^p(D)$ dans $L^p(D)$, de norme $\|U\|\leqslant C_1^{1-\sigma/p}\,C_2^{\sigma/p}$ et, par suite, tout ce qui a été dit sur l'équation (1) est valable pour l'équation (11).

En particulier, si le noyau $K(s, t)$ est de type potentiel, c'est-à-dire si

$$K(s,\ t)=\frac{B(s,\ t)}{|s-t|^m},$$

où $B(s, t)$ est une fonction continue bornée pour $s\neq t$, et $m<n$, alors les conditions indiquées plus haut seront remplies (théorème XI.3.6, tome 1).

Si, de plus, $B(s, t)$ est une fonction continue et

$$p>\frac{n}{n-m},$$

alors d'après le théorème XI.3.7, tome 1, l'opérateur $U$ sera un opérateur de $L^p(D)$ dans $C(D)$, si bien que dans ce cas les séries entières qui représentent les solutions $x^*(s)$ au voisinage d'une valeur caractéristique ou au voisinage du point 0 seront uniformément convergentes.

Le détail des énoncés est laissé au soin du lecteur.

## § 7. Sous-espaces invariants d'un opérateur. Le problème d'approximation

Le carrefour des années 1960-1970 a été marqué par d'importantes découvertes en théorie des $B$-espaces. L'une d'elle est la caractérisation des espaces hilbertiens donnée par Lindenstrauss et Tzafriri (cf. théorème V.3.3, tome 1). Nous nous arrêtons ici sur deux résultats importants. (La théorie des $B$-espaces dans son état actuel est accessible dans l'ouvrage de Lindenstrauss et Tzafriri.)

**7.1.** Soient X un $B$-espace, $U$ un opérateur linéaire continu dans X. Un ensemble linéaire fermé $E \subset$ X est un sous-espace *invariant* de l'opérateur $U$ si $U(E) \subset E$. Il est évident que $\{0\}$ et X sont des sous-espaces invariants. Un sous-espace invariant $E$ est *non trivial* si $E \neq$ X, $\{0\}$. Si un opérateur possède un vecteur propre, c'est que son sous-espace invariant est de dimension un. Cependant, même un opérateur compact peut ne posséder aucun vecteur propre.

Exemple: l'opérateur intégral de type Volterra (cf. 6.2). Mais cet opérateur possède plusieurs sous-espaces invariants. Par exemple, pour tout $a \in ]0, 1[$ l'ensemble de toutes les fonctions $x \in$ C $[0, 1]$ nulles sur l'intervalle $[0, a]$ est un sous-espace invariant.

En 1935 J. von Neumann a prouvé que tout opérateur non nul compact dans un espace hilbertien possède un sous-espace invariant non trivial *). Dernièrement, V. Lomonossov [1] a donné une démonstration simple d'un fait plus général (nous indiquons la variante de la démonstration proposée par Hilden).

Soit $U$ un opérateur compact dans un $B$-espace X de dimension infinie. Désignons par $\mathfrak{A}$ l'ensemble des opérateurs commutant avec $U$, c'est-à-dire l'ensemble de tous les $T \in B$ (X, X) tels que $TU = UT$. Il est évident que $\mathfrak{A}$ est un sous-espace vectoriel de $B$ (X, X), stable pour la multiplication des opérateurs.

THÉORÈME 1 (Lomonossov). *Il existe un sous-espace E dans* X, *qui est un sous-espace invariant non trivial pour tous les opérateurs* $T \in \mathfrak{A}$.

DÉMONSTRATION. Comme $U \neq 0$, on peut admettre que $\| U \| =$
$= 1$. Il existe en outre un élément $x_0 \in$ X tel que

$$\| x_0 \| > 1 \quad \text{et} \quad \| U(x_0) \| > 1. \tag{1}$$

Soit $B$ une boule unité ouverte de centre $x_0$. L'opérateur $U$ étant compact, $K = \overline{U(B)}$ l'est également. Mettons tout $x \in B$ sous la forme $x = x_0 + z$, où $\| z \| < 1$. Si $y = U(x)$, alors $y = U(x_0) +$
$+ U(z)$ et $\| U(z) \| \leqslant 1$. Donc

$$\| y \| \geqslant \| U(x_0) \| - \|U(z)\| \geqslant \| U(x_0) \| - 1 = \delta > 0,$$

d'où

$$\| y \| \geqslant \delta > 0, \quad \forall y \in K, \tag{2}$$

où $\delta$ ne dépend pas de $y$. Distinguons deux cas.

a) L'opérateur $U$ possède une valeur propre non nulle $\lambda$. Désignons par $N_\lambda$ le sous-espace propre associé à $\lambda$. Montrons que $N_\lambda$ est un sous-espace invariant pour tout $T \in \mathfrak{A}$ (il est évident que

*) Ce résultat de Neumann a été publié dans l'article d'Aronszain et Smith [1] où il est généralisé aux $B$-espaces.

$N_\lambda \neq \mathbf{X}$, puisque $U$ est compact). Si $x \in N_\lambda$, alors

$$U(T(x)) = T(U(x)) = T(\lambda x) = \lambda T(x),$$

d'où $T(x) \in N_\lambda$.

b) L'opérateur $U$ ne possède pas de valeurs propres. Le spectre $\sigma(U) = \{0\}$. Donc, le rayon spectral de l'opérateur $U$ est nul, et, d'après 4.1, pour tout $\lambda$

$$\| \lambda^n U^{n+1} \| \xrightarrow[n\to\infty]{} 0. \tag{3}$$

Admettons qu'il n'existe pas de sous-espace non trivial, invariant pour tous les $T \in \mathfrak{A}$. Pour un $x \neq 0$ de $\mathbf{X}$ considérons le sous-espace $\mathbf{L}_x$, adhérence de $\mathbf{L}'_x = \{T(x) : T \in \mathfrak{A}\}$. $\mathfrak{A}$ étant un espace vectoriel, $\mathbf{L}'_x$ l'est aussi. Si $y \in \mathbf{L}'_x$, alors $y = T_0(x)$, $T_0 \in \mathfrak{A}$. Pour tout $T \in \mathfrak{A}$ on a $TT_0 \in \mathfrak{A}$, d'où $T(y) = (TT_0)(y) \in \mathbf{L}'_x$. Ceci et la continuité des opérateurs de $\mathfrak{A}$ entraînent que le sous-espace $\mathbf{L}_x$ est invariant pour tout opérateur de $\mathfrak{A}$.

Comme l'opérateur identique $I$ appartient à $\mathfrak{A}$, $x \in \mathbf{L}'_x$, d'où $\mathbf{L}_x \neq \{0\}$. $\mathbf{L}_x = \mathbf{X}$ pour tout $x \neq 0$, puisque nous avons admis qu'il n'existait pas de sous-espace invariant non trivial, donc $\mathbf{L}'_x \cap B$ n'est pas vide.

Par suite, pour tout $x \neq 0$ il existe $T \in \mathfrak{A}$ tel que $T(x) \in B$, d'où

$$\bigcup_{T\in\mathfrak{A}} T^{-1}(B) = \mathbf{X} \setminus \{0\}.$$

D'après (2), $K \subset \mathbf{X} \setminus \{0\}$, d'où $\{T^{-1}(B) : T \in \mathfrak{A}\}$ est un recouvrement ouvert du compact $K$. Extrayons de $K$ un recouvrement fini:

$$K \subset \bigcup_{i=1}^{m} T_i^{-1}(B).$$

Considérons le point $x_0$ choisi en début de démonstration (on rappelle que $x_0$ vérifie (1)). Comme $U(x_0) \in U(B) \subset K$, il existe $i_1 \in [1, m]$ tel que $U(x_0) \in T_{i_1}^{-1}(B)$. Alors

$$x_1 = T_{i_1}(U(x_0)) \in B.$$

De façon analogue, il existe $i_2 \in [1, m]$ tel que

$$x_2 = T_{i_2}(U(x_1)) \in B,$$

etc. On obtient une suite $\{x_n\} \subset B$ telle que

$$x_n = T_{i_n} U T_{i_{n-1}} U T_{i_{n-2}} \ldots T_{i_1} U(x_0).$$

Comme les opérateurs $T_i$ commutent avec $U$, il vient

$$x_n = T_{i_n} T_{i_{n-1}} T_{i_{n-2}} \ldots T_{i_1} U^n(x_0).$$

Appliquons l'opérateur $U$ à $x_n$:

$$z_n = U(x_n) = T_{i_n} T_{i_{n-1}} T_{i_{n-2}} \ldots T_{i_1} U^{n+1}(x_0) \in K.$$

En vertu de (2)

$$\| z_n \| \geqslant \delta > 0 \quad (n \in \mathbb{N}). \tag{4}$$

Estimons maintenant $\| z_n \|$ d'une autre manière. Posons $\lambda = \max\{\| T_i \| : 1 \leqslant i \leqslant n\}$. On a

$$\| z_n \| \leqslant \prod_{j=1}^{n} \| T_{i_j} \| \| U^{n+1} \| \| x_0 \| \leqslant \| \lambda^n U^{n+1} \| \| x_0 \|. \tag{5}$$

D'après (3), il résulte que le second membre de (5) tend vers 0 pour $n \to \infty$, ce qui contredit (4). Cette contradiction prouve le théorème. De ce théorème il suit, en particulier, que l'opérateur $U$ possède un sous-espace invariant non trivial, c'est-à-dire on retrouve l'ancien résultat de Neumann.

Nous avons vu qu'il n'est pas facile de déterminer les sous-espaces invariants non triviaux d'une certaine classe d'opérateurs. Tout aussi compliqué est le problème inverse qui consiste à trouver un opérateur linéaire continu agissant dans un espace de Banach séparable sans sous-espaces invariants non triviaux.

**7.2.** Nous avons vu dans ce chapitre que les opérateurs compacts s'apparentent par leurs propriétés aux opérateurs de dimension finie. Ceci a suggéré à A. Grothendieck [2] la définition suivante.

On dit qu'un $B$-espace **X** possède la *propriété d'approximation* si tout opérateur compact $U$ d'un $B$-espace **Y** dans **X** est la limite d'une suite d'opérateurs de dimension finie pour la norme de l'espace $B$ (**Y**, **X**). Ceci revient à dire que pour tout compact $K$ dans **X** et tout nombre $\varepsilon > 0$ il existe un opérateur de dimension finie $T \in B(\mathbf{X}, \mathbf{X})$ tel que

$$\| Tx - x \| < \varepsilon \quad (x \in K).$$

Grothendieck s'est demandé si tout $B$-espace possédait la propriété d'approximation. Il a prouvé que la résolution de ce problème, dit problème d'approximation, passait par la réalisation de conditions assez concrètes. Par exemple, par la suivante. Si $K(s, t)$ est une fonction continue sur le carré $[0, 1; 0, 1]$ et

$$\int_0^1 K(s, t) K(t, u)\, dt = 0, \quad \forall s, \forall u \in [0, 1],$$

alors

$$\int_0^1 K(t, t)\, dt = 0.$$

Le problème de la base, qui consiste en ce qui suit, a été posé antérieurement (voir Banach).

On dit qu'une suite $\{x_n\}$ d'éléments d'un $B$-espace X est une *base* *) dans X si chaque élément $x \in X$ peut être représenté d'une seule manière par la série

$$x = \sum_{n=1}^{\infty} \lambda_n x_n$$

convergeant pour la norme de l'espace X.

Il est évident qu'un espace muni d'une base est nécessairement séparable. Tout système orthonormal complet dans un espace hilbertien (cf. IV.5.8, tome 1) est une base. Il y a des bases dans les espaces $L^p(0, 1)$ et $l^p$, $1 \leqslant p < \infty$, $C[0, 1]$. Dans $l^p$ par exemple, la suite des vecteurs unités, c'est-à-dire des éléments $x_n = \{\xi_k\}_{k=1}^{\infty}$, où $\xi_n = 1$, $\xi_k = 0$ $(k \neq n)$, forme manifestement une base.

Le problème de la base est le problème qui consiste à établir si tout $B$-espace séparable possède une base.

On démontre sans peine que tout espace muni d'une base possède la propriété d'approximation.

Les deux problèmes mentionnés furent longtemps en suspens. Ce n'est qu'en 1972 que le mathématicien suédois Enflo produisit un exemple remarquable donnant une réponse négative à la fois aux problèmes d'approximation et de la base (cf. Enflo [1]). Plus exactement, il construisit un exemple de $B$-espace réflexif séparable sans propriété d'approximation, donc sans base. Pour plus de détails sur ce sujet voir l'article de Peltchinski et Figuel [1].

En 1975, A. Chankovski a construit un exemple d'e.f. de Banach réflexif séparable sans propriété d'approximation.

---

*) La notion de base est due à Schauder [1]. Pour de plus amples détails sur les bases voir Singer.

CHAPITRE XIV

# THÉORIE GÉNÉRALE DES MÉTHODES D'APPROXIMATION

Un grand nombre de méthodes fondées sur des principes différents ont été proposées et sont pratiquement utilisées pour la résolution approchée de problèmes d'analyse mathématique: équations différentielles et intégrales, problèmes aux limites de physique mathématique, application conforme, etc. Ainsi, pour les problèmes aux limites on se sert des méthodes variationnelles ou semblables aux variationnelles (méthodes de Ritz, de Galerkin, des moments, de réduction aux équations différentielles ordinaires), des méthodes aux différences, des méthodes d'interpolation. Pour juger de l'efficacité et du bien-fondé de l'application de ces méthodes, il est nécessaire d'en entreprendre une étude théorique dans laquelle se posent dans un ordre croissant de précision et de complexité les trois problèmes suivants:

a) établissement de la réalisabilité et de la convergence de l'algorithme;

b) étude de la vitesse de convergence;

c) estimation effective de l'erreur.

Ces problèmes ont été différemment résolus pour chaque classe d'équations et pour chaque méthode. La résolution a soulevé souvent de grosses difficultés, dont certaines n'ont pu être tournées à ce jour.

Il importe de regrouper ces recherches et de bâtir une théorie unique des méthodes d'approximation. Les idées de l'analyse fonctionnelle fournissent une approche naturelle de la résolution de ce problème.

Dans ce chapitre nous considérons la construction de la théorie des méthodes d'approximations pour une classe d'équations linéaires. Dans cette théorie, les équations exacte et approchée seront envisagées dans des espaces différents qui, certes, sont reliés d'une certaine façon, comme c'est le cas dans de nombreuses méthodes concrètes, par exemple dans le changement approché d'une équation intégrale par un système fini d'équations algébriques linéaires. Du reste, on montrera qu'il sera toujours possible de se ramener au cas où l'espace

approximant est un sous-espace de l'espace dans lequel est donnée l'équation exacte.

Dans la suite on prouvera des théorèmes de double nature: des théorèmes qui permettent sur la base des données sur l'équation exacte d'établir la résolubilité de l'équation « approchée » et la convergence de la solution approchée vers la solution analytique, et, au contraire, des théorèmes qui, sur la base des résultats de la résolution approchée, nous renseignent sur la solubilité de l'équation exacte et de l'écart entre les deux solutions.

Les résultats de ce chapitre sous des hypothèses légèrement différentes appartiennent à L. Kantorovitch (cf. Kantorovitch [9]). Certains théorèmes du § 1 ont été précisés par G. Akilov. D'autres constructions de la théorie des méthodes d'approximation utilisant l'arsenal de l'analyse fonctionnelle sont accessibles dans les travaux de S. Mikhline, M. Krasnosselski, N. Polski, G. Vaïnikko, M. Gavourine et autres. Cf. Vaïnikko; Gavourine; Krasnosselski et autres; Krasnosselski [1]; Mikhline [II].

## § 1. Théorie générale pour les équations de seconde espèce

**1.1.** Soient $\mathbf{X}$ un espace normé, $\tilde{\mathbf{X}}$ un sous-espace complet *) de $\mathbf{X}$. Supposons qu'il existe un opérateur linéaire continu $P$ de $\mathbf{X}$ dans $\tilde{\mathbf{X}}$, c'est-à-dire tel que

$$P(\mathbf{X}) = \tilde{\mathbf{X}}, \quad P^2 = P.$$

Il est clair que l'opérateur $P$ laisse invariants les éléments de $\tilde{\mathbf{X}}$.

Supposons par exemple que $\mathbf{X} = \mathbf{C}[a, b]$ et $\tilde{\mathbf{X}}$ est l'ensemble des polynômes de degré $\leqslant (n-1)$. L'opérateur $P$ associe à une fonction continue $x \in \mathbf{C}[a, b]$ son polynôme d'interpolation construit sur un système de nœuds $t_1, t_2, \ldots, t_n$ donné à l'avance.

Considérons par ailleurs deux équations, l'une

$$Kx \equiv x - \lambda Hx = y \tag{1}$$

dans l'espace $\mathbf{X}$, l'autre

$$\tilde{K}\tilde{x} \equiv \tilde{x} - \lambda\tilde{H}\tilde{x} = Py \tag{2}$$

dans l'espace $\tilde{\mathbf{X}}$. Ici $H$ est un opérateur linéaire continu dans $\mathbf{X}$, $\tilde{H}$ un opérateur linéaire continu dans $\tilde{\mathbf{X}}$. L'équation (1) est dite *exacte*, l'équation (2) *approchée* (associée à l'équation exacte (1)).

Les espaces $\mathbf{X}$ et $\tilde{\mathbf{X}}$ et les opérateurs $H$ et $\tilde{H}$ seront dans la suite reliés par les conditions suivantes.

I. (Condition de proximité des opérateurs $H$ et $\tilde{H}$.) Pour tout $\tilde{x} \in \tilde{\mathbf{X}}$

$$\| PH\tilde{x} - \tilde{H}\tilde{x} \| \leqslant \eta \| \tilde{x} \|.$$

*) $\mathbf{X}$ n'est pas supposé être complet.

Cette relation est équivalente à la suivante:

$$\| PK\tilde{x} - \tilde{K}\tilde{x} \| \leqslant |\lambda| \, \eta \, \| \tilde{x} \| \quad (\tilde{x} \in \tilde{X}).$$

II. (Condition de bonne approximation des éléments de la forme $Hx$ par des éléments de $\tilde{X}$.) Pour tout $x \in X$ il existe $\tilde{x} \in \tilde{X}$ tel que

$$\| Hx - \tilde{x} \| \leqslant \eta_1 \| x \|.$$

III. (Condition de bonne approximation du second membre de l'équation exacte.) Il existe un élément $\tilde{y} \in \tilde{X}$ tel que

$$\| y - \tilde{y} \| \leqslant \eta_2 \| y \|.$$

Contrairement aux conditions précédentes, ici $\eta_2$ dépend généralement de $y$.

1.2. Pour les équations vérifiant les conditions énumérées (ou certaines d'entre elles) nous allons prouver des théorèmes liant l'équation exacte à l'équation approchée.

Théorème 1 (de solubilité de l'équation approchée). *Supposons que les conditions* I *et* II *sont réunies et que l'opérateur $K$ possède un inverse continu. Si*

$$q = |\lambda| \, [\eta + \| I - P \| \, \eta_1] \, \| K^{-1} \| < 1, \tag{3}$$

*l'opérateur $\tilde{K}$ possède aussi un inverse continu $\tilde{K}^{-1}$. De plus*

$$\| \tilde{K}^{-1} \| \leqslant \frac{\| K^{-1} \|}{1 - q}. \tag{4}$$

Démonstration. Supposons tout d'abord que l'espace $X$ est complet. Considérons l'opérateur $K_1 = I - \lambda PH$. Pour un élément quelconque $x$ de $X$, construisons $\tilde{x} \in \tilde{X}$ suivant la condition II, soit

$$\| Hx - \tilde{x} \| \leqslant \eta_1 \| x \|.$$

Alors

$$\| (K - K_1) x \| = |\lambda| \, \| Hx - PHx \| = |\lambda| \, \| Hx - \tilde{x} + P\tilde{x} - PH\tilde{x} \| =$$
$$= |\lambda| \, \| (I - P)(Hx - \tilde{x}) \| \leqslant |\lambda| \, \| I - P \| \, \eta_1 \, \| x \|.$$

L'élément $x$ étant arbitraire, on a prouvé que

$$\| K - K_1 \| \leqslant |\lambda| \, \| I - P \| \, \eta_1.$$

L'opérateur $K_1$ peut être mis sous la forme

$$K_1 = K(I - K^{-1}(K - K_1)), \tag{5}$$

de plus pour l'opérateur $K^{-1}(K - K_1)$ on a la majoration

$$\| K^{-1}(K - K_1) \| \leqslant |\lambda| \, \| I - P \| \, \eta_1 \, \| K^{-1} \| \leqslant q < 1.$$

Le théorème de Banach (théorème V.4.3, tome 1) affirme l'existence de l'opérateur inverse $(I - K^{-1}(K - K_1))^{-1}$ et

$$\| (I - K^{-1}(K - K_1))^{-1} \| \leqslant \frac{1}{1 - |\lambda| \, \| I - P \| \, \eta_1 \, \| K^{-1} \|}.$$

De la représentation (5) il résulte maintenant que l'opérateur $K_1$ est inversible: $K_1^{-1} = (I - K^{-1}(K - K_1))^{-1} K^{-1}$, donc

$$\| K_1^{-1} \| \leqslant \frac{\| K^{-1} \|}{1 - |\lambda| \, \| I - P \| \, \eta_1 \, \| K^{-1} \|} \tag{6}$$

Dans l'espace $\widetilde{X}$, considérons l'opérateur $\widetilde{K}_1 \tilde{x} = \tilde{x} - \lambda PH\tilde{x}$. Il est évident que $\widetilde{K}_1 \tilde{x} = K_1 \tilde{x}$ pour tout $\tilde{x} \in \widetilde{X}$. L'opérateur $K_1$ jouit encore de la propriété suivante: si $\tilde{y} \in \widetilde{X}$, alors $K_1^{-1}\tilde{y} \in \widetilde{X}$. En effet, si $x' = K_1^{-1}\tilde{y}$, alors $x' - PHx' = \tilde{y}$, $x' = \tilde{y} + PHx' \in \widetilde{X}$. Donc, l'opérateur $\widetilde{K}_1$ admet un inverse continu confondu avec $K_1^{-1}$ sur $\widetilde{X}$ et

$$\| \widetilde{K}_1^{-1} \| \leqslant \| K_1^{-1} \|. \tag{7}$$

Majorons maintenant la différence $\widetilde{K} - \widetilde{K}_1$. D'après la condition 1 pour tout $\tilde{x} \in X$

$$\| \widetilde{K}\tilde{x} - \widetilde{K}_1 \tilde{x} \| = |\lambda| \, \| PH\tilde{x} - \widetilde{H}\tilde{x} \| \leqslant |\lambda| \, \eta \, \| \tilde{x} \|, \tag{8}$$

et par conséquent

$$\| \widetilde{K} - \widetilde{K}_1 \| \leqslant |\lambda| \, \eta. \tag{9}$$

Donc, d'après (7), (6) et (9)

$$\| \widetilde{K}_1^{-1} \| \, \| \widetilde{K} - \widetilde{K}_1 \| \leqslant \frac{\| K^{-1} \| \, |\lambda| \, \eta}{1 - |\lambda| \, \| I - P \| \, \eta_1 \, \| K^{-1} \|} =$$

$$= 1 - \frac{1 - q}{1 - |\lambda| \, \| I - P \| \, \eta_1 \, \| K^{-1} \|} < 1. \tag{10}$$

En mettant l'opérateur $\widetilde{K}$ sous la forme $\widetilde{K} = \widetilde{K}_1 (I - \widetilde{K}_1^{-1} (\widetilde{K}_1 - K))$ et en appliquant encore le théorème de Banach, on prouve l'existence de l'opérateur inverse $\widetilde{K}^{-1}$ et la majoration

$$\| \widetilde{K}^{-1} \| \leqslant \frac{\| \widetilde{K}_1^{-1} \|}{1 - \| \widetilde{K}_1^{-1} \| \, \| \widetilde{K} - \widetilde{K}_1 \|} \leqslant$$

$$\leqslant \frac{(1 - |\lambda| \, \| I - P \| \, \eta_1 \, \| K^{-1} \|) \, \| K_1^{-1} \|}{1 - q} \leqslant \frac{\| K^{-1} \|}{1 - q}. \tag{11}$$

Ce qui achève la démonstration du théorème dans le cas d'un espace X complet.

Le cas d'un espace X non complet se ramène sans peine au précédent. En effet, si l'on introduit la complétion de l'espace X, les fermetures (cf. théorème V.8.2, tome 1) des opérateurs $K$ et $K^{-1}$ seront respectivement inverses, la fermeture de l'opérateur $H$ sera

également liée à l'opérateur $\widetilde{H}$ par les conditions I et II avec la même constante $\eta$ et une nouvelle constante $\eta_1$ aussi proche que l'on veut de l'ancienne.

**1.3.** Il est naturel de considérer la solution de l'équation approchée (2) comme une solution approchée de l'équation exacte (1). La majoration de l'erreur de cette solution approchée est établie par le théorème suivant.

THÉORÈME 2 (de majoration de l'erreur de la solution approchée). *Si les conditions* I, II *et* III *sont remplies, l'opérateur inverse* $\widetilde{K}^{-1}$ *existe (notamment si sont réalisées les conditions du théorème* 1) *et l'équation* (1) *admet la solution* $x^*$, *alors*

$$\| x^* - \widetilde{x}^* \| \leqslant p \| x^* \|, \tag{12}$$

*où* $\widetilde{x}^*$ *est solution de l'équation* (2), *et*

$$p = 2 |\lambda| \eta \| \widetilde{K}^{-1} \| + (\eta_1 |\lambda| + \eta_2 \| K \|)(1 + \| \widetilde{K}^{-1} PK \|). \tag{13}$$

DÉMONSTRATION. Prouvons tout d'abord que les conditions du théorème permettent d'approcher l'élément $x^*$ par un élément de $\widetilde{X}$ avec une précision de l'ordre de $\eta_1 + \eta_2$. Plus exactement prouvons l'existence d'un élément $\widetilde{x} \in \widetilde{X}$ tel que

$$\| x^* - \widetilde{x} \| \leqslant \varepsilon \| x^* \|, \tag{14}$$

où

$$\varepsilon = \min [1, \ \eta_1 |\lambda| + \eta_2 \| K \|]. \tag{15}$$

En effet, en prenant $\widetilde{y} \in \widetilde{X}$ et $\widetilde{z} \in \widetilde{X}$ d'après II et III tels que

$$\| Hx^* - \widetilde{z} \| \leqslant \eta_1 \| x^* \|, \quad \| y - \widetilde{y} \| \leqslant \eta_2 \| y \| \leqslant \eta_2 \| K \| \| x^* \|,$$

et en posant $\widetilde{x} = \lambda \widetilde{z} + \widetilde{y}$, on aura

$$\| x^* - \widetilde{x} \| = \| \lambda H x^* + y - (\lambda \widetilde{z} + \widetilde{y}) \| \leqslant (\eta_1 |\lambda| + \eta_2 \| K \|) \| x^* \|.$$

D'autre part (14) est réalisée pour $\varepsilon = 1$ si l'on pose $\widetilde{x} = 0$. D'où il suit que pour $\varepsilon$ on peut prendre (15).

Passons maintenant à la démonstration de l'inégalité (12).

En désignant $\widetilde{x}_0 = \widetilde{K}^{-1} PK \widetilde{x}$, on aura

$$\| x^* - \widetilde{x}^* \| \leqslant \| x^* - \widetilde{x} \| + \| \widetilde{x} - \widetilde{x}_0 \| + \| \widetilde{x}_0 - \widetilde{x}^* \|. \tag{16}$$

Evaluons chaque terme séparément. Le premier terme a été majoré dans (14).

Le second se majore comme suit : d'après la condition I

$$\| \widetilde{x} - \widetilde{x}_0 \| = \| \widetilde{K}^{-1} \widetilde{K} \widetilde{x} - \widetilde{K}^{-1} PK \widetilde{x} \| \leqslant \| \widetilde{K}^{-1} \| \| \widetilde{K} \widetilde{x} - PK \widetilde{x} \| \leqslant$$
$$\leqslant |\lambda| \eta \| \widetilde{K}^{-1} \| \| \widetilde{x} \|. \tag{17}$$

Mais d'après (14)

$$\|\tilde{x}\| \leqslant \|x^*\| + \|x^* - \tilde{x}\| \leqslant (1+\varepsilon)\|x^*\|.$$

En portant ceci dans (17), on obtient en définitive

$$\|\tilde{x} - \tilde{x}_0\| \leqslant |\lambda|\, \eta\, (1+\varepsilon)\, \|\tilde{K}^{-1}\|\, \|x^*\|. \tag{18}$$

Majorons le dernier terme de (16) en tenant compte de ce que $\tilde{x}^*$ est solution de l'équation (2), donc que

$$\tilde{x}^* = \tilde{K}^{-1}PKx^*.$$

D'après (14)

$$\|\tilde{x}_0 - \tilde{x}^*\| = \|\tilde{K}^{-1}PK\tilde{x} - \tilde{K}^{-1}PKx^*\| \leqslant$$
$$\leqslant \|\tilde{K}^{-1}PK\|\, \|\tilde{x} - x^*\| \leqslant \varepsilon\, \|\tilde{K}^{-1}PK\|\, \|x^*\|.$$

En combinant ceci avec (14) et (18), on arrive à la majoration

$$\|x^* - \tilde{x}^*\| \leqslant [|\lambda|\, \eta\, (1+\varepsilon)\, \|\tilde{K}^{-1}\| + \varepsilon\, (1 + \|\tilde{K}^{-1}PK\|)]\, \|x^*\|. \tag{19}$$

Reste à remarquer que $1+\varepsilon \leqslant 2$ et $\varepsilon \leqslant \eta_1\, |\lambda| + \eta_2\, \|K\|$. C.q.f.d.

Remarque 1. Parfois on peut établir directement, sans passer par les conditions II et III, la possibilité d'approcher $x^*$ par un élément de $\tilde{X}$. Dans ce cas on peut encore se servir de l'inégalité (19), et l'on peut bien sûr se dispenser des conditions II et III.

Il est utile de remarquer que dans ce cas il est question d'un élément $x^*$ concret dont $\varepsilon$ peut dépendre, alors que dans la condition II on a supposé que $\eta_1$ ne dépendait pas de $x$.

Remarque 2. Grâce au théorème 2 il est facile de majorer l'écart entre la solution approchée et la solution exacte sans recourir aux données relatives à la solution exacte. Plus exactement, si l'on se place dans les conditions du théorème 2 et si de plus $p < 1$, alors

$$\|x^* - \tilde{x}^*\| \leqslant \frac{p}{1-p}\, \|\tilde{x}^*\|. \tag{20}$$

En effet

$$\|x^*\| \leqslant \|\tilde{x}^*\| + \|x^* - \tilde{x}^*\|.$$

En tenant compte de ceci dans l'inégalité (12), on obtient

$$\|x^* - \tilde{x}^*\| \leqslant p\, \|\tilde{x}^*\| + p\, \|x^* - \tilde{x}^*\|,$$

d'où l'on déduit sans peine (20).

1.4. Soient une suite d'équations approchées et les solutions approchées obtenues à l'aide de ces équations. Dans ce cas l'espace $\tilde{X}$, les opérateurs $\tilde{H}$ $(\tilde{K})$, $P$ et les constantes $\eta$, $\eta_1$, $\eta_2$, $q$, $p$, $\varepsilon$, . . . dépendent de l'indice $n$ que nous omettrons pour alléger l'écriture.

Le théorème suivant exprime les conditions de convergence de la suite $\{\tilde{x}_n^*\}$ de solutions approchées vers la solution exacte.

Théorème 3. *Si*

1) *l'opérateur* $K$ *admet un inverse continu;*

2) *pour chaque* $n = 1, 2, \ldots$ *sont réalisées les conditions* I, II, III *et de plus*

$$\lim_{n\to\infty} \eta = 0, \quad \lim_{n\to\infty} \eta_1 \| P \| = 0, \quad \lim_{n\to\infty} \eta_2 \| P \| = 0, \tag{21}$$

*alors pour* $n$ *assez grand les équations approchées admettent chacune une solution et la suite de ces solutions approchées converge vers la solution exacte:*

$$\lim_{n\to\infty} \| x^* - \tilde{x}_n^* \| = 0.$$

*Plus exactement*

$$\| x^* - \tilde{x}_n^* \| \leqslant Q_0 \eta + Q_1 \eta_1 \| P \| + Q_2 \eta_2 \| P \|, \tag{22}$$

*où* $Q_0$, $Q_1$, $Q_2$ *sont des constantes.*

Démonstration. Comme $\| P \| \geqslant 1$, de (21) il suit que $\lim\limits_{n\to\infty} \eta_1 = 0$, donc, pour $n$ assez grand, on aura $q < 1/2$ dans le théorème 1. Par suite, pour les $n$ indiqués il existe l'opérateur continu $\tilde{K}^{-1}$ et de plus

$$\| \tilde{K}^{-1} \| \leqslant \frac{\| K^{-1} \|}{1-q} < 2 \| K^{-1} \|.$$

D'où il suit que $\| \tilde{K}^{-1} \|$ est bornée indépendamment de $n$. En tenant compte de ceci et grâce à la majoration (12) du théorème 2 on obtient (22).

Remarque. Si dans la démonstration on se sert de la majoration (19) au lieu de (12), on obtient

$$\| x^* - \tilde{x}_n^* \| \leqslant Q' \eta + Q' \varepsilon \| P \| \tag{23}$$

et les conditions de convergence (21) sont remplacées par

$$\lim_{n\to\infty} \eta = 0, \quad \lim_{n\to\infty} \varepsilon \| P \| = 0. \tag{24}$$

La condition III est superflue (s'agissant de la condition II, elle est nécessaire, car elle est utilisée dans le théorème 1; du reste, la condition $\eta_1 \| P \| \to 0$ peut être remplacée par $q \leqslant q_0 < 1$ pour $n$ assez grand).

1.5. Le théorème 1 permet d'établir la solubilité de l'équation approchée en fonction de celle de l'équation exacte. Le théorème suivant permet de tirer la conclusion inverse.

Théorème 4. *Si l'opérateur* $\widetilde{K}$ *admet un inverse continu, les conditions* I *et* II *sont remplies et*

$$r = |\lambda|\, \eta\, (1 + |\lambda|\, \eta_1)\, \| \widetilde{K}^{-1} \| + |\lambda|\, \eta_1\, (1 + \| \widetilde{K}^{-1}PK \|) < 1, \qquad (25)$$

*l'opérateur* $K$ *admet un inverse à gauche continu dont la norme est majorée par*

$$\| K^{-1} \| \leqslant \frac{1 + \| \widetilde{K}^{-1}P \| + |\lambda|\, \eta\, \| \widetilde{K}^{-1} \| + \| \widetilde{K}^{-1}PK \|}{1 - r}. \qquad (26)$$

Démonstration. Soit $x^*$ un élément quelconque non nul de l'espace X. Il est manifestement solution de l'équation

$$Kx = Kx^*.$$

D'après la condition II il existe un élément $\widetilde{x} \in X$ tel que

$$\| Hx^* - \widetilde{x} \| \leqslant \eta_1\, \| x^* \|.$$

Comme $x^* = \lambda Hx^* + Kx^*$, il vient

$$\| x^* - \lambda \widetilde{x} \| \leqslant |\lambda|\, \eta_1\, \| x^* \| + \| Kx^* \| = \left( |\lambda|\, \eta_1 + \frac{\| Kx^* \|}{\| x^* \|} \right) \| x^* \|.$$

Cette inégalité permet d'utiliser le théorème 2 (sous la forme de la remarque 1 qui le suit) avec la valeur

$$\varepsilon = |\lambda|\, \eta_1 + \frac{\| Kx^* \|}{\| x^* \|}. \qquad (27)$$

L'équation approchée est ici

$$\widetilde{K}\widetilde{x} = PKx^*$$

et sa solution est $\widetilde{x}^* = \widetilde{K}^{-1}PKx^*$. En portant dans (19) la valeur de $\varepsilon$ tirée de (27), on obtient

$$\| x^* - \widetilde{K}^{-1}PKx^* \| \leqslant \Big[ |\lambda|\, \eta \left( 1 + |\lambda|\, \eta_1 + \frac{\| Kx^* \|}{\| x^* \|} \right) \| \widetilde{K}^{-1} \| + $$
$$+ \left( |\lambda|\, \eta_1 + \frac{\| Kx^* \|}{\| x^* \|} \right) (1 + \| \widetilde{K}^{-1}PK \|) \Big] \| x^* \| =$$
$$= [ |\lambda|\, \eta\, (1 + |\lambda|\, \eta_1)\, \| \widetilde{K}^{-1} \| + |\lambda|\, \eta_1\, (1 + \| \widetilde{K}^{-1}PK \|)]\, \| x^* \| +$$
$$+ [ |\lambda|\, \eta\, \| \widetilde{K}^{-1} \| + 1 + \| \widetilde{K}^{-1}PK \| ]\, \| Kx^* \|.$$

D'où

$$\| x^* \| \leqslant \| x^* - \widetilde{K}^{-1}PKx^* \| + \| \widetilde{K}^{-1}P \|\, \| Kx^* \| \leqslant$$
$$\leqslant r\, \| x^* \| + [1 + \| \widetilde{K}^{-1}P \| + |\lambda|\, \eta\, \| \widetilde{K}^{-1} \| + \| \widetilde{K}^{-1}PK \| ]\, \| Kx^* \|$$

ou

$$\| Kx^* \| \geqslant \frac{1 - r}{1 + \| K^{-1}P \| + |\lambda|\, \eta\, \| \widetilde{K}^{-1} \| + \| \widetilde{K}^{-1}PK \|}\, \| x^* \| = Q\, \| x^* \|.$$

$Q > 0$, puisque $r < 1$, ce qui prouve le théorème d'après ce qui a été dit dans V.4.4, tome 1.

COROLLAIRE. *Si l'opérateur $K$ vérifie la condition:*

(A) *Si l'équation* (1) *admet une solution unique, elle en admet une pour tout second membre;*

*alors dans les conditions du théorème existe l'opérateur inverse continu* $K^{-1}$.

1.6. Si la condition III est réalisée avec une grande constante $\eta_2$, c'est-à-dire si le second membre de l'équation (1) n'admet pas une bonne approximation par les éléments de $\widetilde{X}$, la solution $x^*$ n'admet pas une bonne approximation et la solution approchée $\widetilde{x}^*$ sera manifestement mauvaise.

Dans ce cas, on a intérêt à « régulariser » le second membre. Effectuons le changement $x = y + z$ dans l'équation (1). La nouvelle équation (en $z$) est

$$Kz = y' \quad (y' = \lambda Hy). \tag{28}$$

Le second membre de cette équation est « régulier » d'après la condition II.

L'équation approchée de (28) sera

$$\widetilde{K}\widetilde{z} = \lambda PHy. \tag{29}$$

Désignons par $\widetilde{z}^*$ la solution de cette équation et considérons comme solution approchée de l'équation (1) l'élément

$$x' = y + \widetilde{z}^*. \tag{30}$$

Comme

$$\| x^* - x' \| = \| z^* - \widetilde{z}^* \|,$$

où $z^*$ est la solution de l'équation (28), l'erreur de la solution approchée (30) se majore à l'aide du théorème 2 (pour $\eta_2$ il convient de prendre $|\lambda|\eta_1$ *)). Cependant, l'étude directe de ce cas permet d'obtenir une meilleure majoration.

THÉORÈME 5. *Si les conditions* I *et* II *sont remplies et si existent l'opérateur inverse continu* $\widetilde{K}^{-1}$ *et la solution* $x^*$ *de l'équation* (1), *alors*

$$\| x^* - x' \| \leqslant p' \| x^* \|, \tag{31}$$

*) Voire même la quantité moindre $|\lambda|\eta_1^*$, pourvu que

$$\| Hy - \widetilde{y} \| \leqslant \eta_1^* \| y \|$$

pour un $\widetilde{y} \in \widetilde{X}$.

Donc, si la deuxième des conditions (21) est réalisée, on peut toujours obtenir la réalisation de la troisième.

*où*

$$p' = |\lambda| \, [\eta_1 (1 + \| \tilde{K}^{-1} P K \|) + |\lambda| \, \eta \, \| \tilde{K}^{-1} \| \, (\| H \| + \eta_1)]. \quad (32)$$

DÉMONSTRATION. Comme

$$z^* = \lambda H z^* + \lambda H y = \lambda H x^*, \quad (33)$$

d'après la condition II on peut trouver un élément $\tilde{z} \in \tilde{X}$ tel que

$$\| z^* - \tilde{z} \| \leqslant |\lambda| \, \eta_1 \, \| x^* \|. \quad (34)$$

Posons

$$\tilde{z}_0 = \tilde{K}^{-1} P K \tilde{z}.$$

Puisque

$$\tilde{K} \tilde{z}^* = P K z^*,$$

alors

$$\tilde{z}^* = \tilde{K}^{-1} P K z^*,$$

d'où

$$\| \tilde{z}^* - \tilde{z}_0 \| \leqslant \| \tilde{K}^{-1} P K \| \, \| z^* - \tilde{z} \| \leqslant |\lambda| \, \eta_1 \, \| \tilde{K}^{-1} P K \| \, \| x^* \|. \quad (35)$$

D'après la condition I, (34) et (33)

$$\| \tilde{z} - \tilde{z}_0 \| \leqslant \| \tilde{K}^{-1} \| \, \| \tilde{K} \tilde{z} - P K \tilde{z} \| \leqslant |\lambda| \, \eta \, \| \tilde{K}^{-1} \| \, \| \tilde{z} \| \leqslant$$
$$\leqslant |\lambda| \, \eta \, \| \tilde{K}^{-1} \| \, [\| z^* \| + \| \tilde{z} - z^* \|] \leqslant$$
$$\leqslant |\lambda| \, \eta \, \| \tilde{K}^{-1} \| \, (|\lambda| \, \| H \| + |\lambda| \, \eta_1) \, \| x^* \|.$$

Mais (34) et (35) entraînent

$$\| x^* - x' \| \leqslant \| z^* - \tilde{z} \| + \| \tilde{z} - \tilde{z}_0 \| + \| \tilde{z}_0 - \tilde{z}^* \| \leqslant [|\lambda| \, \eta_1 +$$
$$+ |\lambda|^2 \eta \, \| \tilde{K}^{-1} \| \, (\| H \| + \eta_1) + |\lambda| \, \eta_1 \, \| \tilde{K}^{-1} P K \|] \, \| x^* \| = p' \, \| x^* \|,$$

c.q.f.d.

On peut de façon analogue améliorer dans ce cas le résultat du théorème 4.

1.7. Etudions la convergence des valeurs caractéristiques de l'opérateur $\tilde{H}$ vers les valeurs caractéristiques de l'opérateur $H$ en nous limitant au cas où ces deux opérateurs sont compacts, donc au cas où leurs ensembles caractéristiques sont discrets.

Tout d'abord, en se plaçant dans les conditions du théorème 3, on peut affirmer que les valeurs caractéristiques de l'opérateur $\tilde{H}$ ne peuvent converger que vers celles de l'opérateur $H$. En effet, considérons dans le plan $\lambda$ un disque $C_0$ de rayon $R$ centré en l'origine des coordonnées et soient $\lambda_1, \lambda_2, \ldots, \lambda_m$ les valeurs caractéristiques de l'opérateur $H$ situées à l'intérieur de $C_0$. Désignons par $D$ le do-

maine fermé obtenu à partir de $C_0$ par élimination des δ-voisinages des points $\lambda_1, \lambda_2, \ldots, \lambda_m$. Dans ce domaine l'opérateur $K_\lambda^{-1}$ ($K_\lambda = I - \lambda H$) existe et $\| K_\lambda^{-1} \|$ est bornée comme fonction de λ. Le théorème 3 nous dit alors que $\| \tilde{K}_\lambda^{-1} \|$ ($\tilde{K}_\lambda = \tilde{I} - \lambda\tilde{H}$) est bornée indépendamment de $n$ assez grand. Donc, sous les conditions indiquées, les valeurs caractéristiques de l'opérateur $\tilde{H}$ ne peuvent se trouver que dans les δ-voisinages mentionnés ou à l'extérieur de $C_0$, d'où il est clair que si pour $n \to \infty$ elles admettent une limite finie, celle-ci sera confondue avec une valeur caractéristique de l'opérateur $H$.

D'autre part, chaque valeur caractéristique de l'opérateur $H$ est limite de valeurs caractéristiques de l'opérateur $\tilde{H}$. En effet, si un $\lambda_k$ n'était pas limite, alors dans un voisinage du point $\lambda_k$, limité par le cercle $C_k$, il n'existerait pas de valeurs caractéristiques de l'opérateur $\tilde{H}$ (pour $n$ assez grand). La fonction $\| K_\lambda^{-1} \|$ est bornée sur le cercle $C_k$, donc la fonction $\| \tilde{K}_\lambda^{-1} \|$ l'est aussi en vertu de (11). Par suite, $\| \tilde{K}_\lambda^{-1} \|$ serait également bornée par le même nombre à l'intérieur de ce cercle. Ce dernier fait résulte du principe du maximum du module, appliqué à la fonction analytique $\psi(\lambda) = f(\tilde{K}_\lambda^{-1}\tilde{x})$ ($f$ est une fonctionnelle arbitraire dans l'espace $\tilde{X}$, $\tilde{x}$ un élément quelconque de $\tilde{X}$). En appliquant le théorème 4 on établirait l'existence de $K_\lambda^{-1}$ dans tout le voisinage, ce qui nous amène à une contradiction.

La convergence des valeurs et éléments propres dans le cadre de la théorie générale des méthodes d'approximation a été étudiée par Troïtskaïa [1] et Vaïnikko (cf. aussi Mikhline [II]).

1.8. Signalons un cas important où il est possible de simplifier l'énoncé des théorèmes prouvés précédemment.

Souvent l'équation approchée (2) est construite d'une manière spéciale. Plus exactement, pour opérateur $\tilde{H}$ on considère l'opérateur $PH$ *); donc l'équation approchée est:

$$\tilde{x} - \lambda PH\tilde{x} = Py \quad (\tilde{K} = PK). \tag{36}$$

La condition I est visiblement réalisée pour $\eta = 0$. Ceci permet de simplifier l'énoncé des théorèmes. Ainsi, dans le théorème 1, il convient de prendre

$$q = |\lambda| \eta_1 \| I - P \| \| K^{-1} \|. \tag{37}$$

Dans le théorème 2

$$p = (\eta_1 |\lambda| + \eta_2 \| K \|)(1 + \| \tilde{K}^{-1} PK \|). \tag{38}$$

---

*) $H$ doit être considéré comme un opérateur de $\tilde{X}$ dans X.

Les conditions (21) du théorème 3 sont remplacées par

$$\lim_{n\to\infty} \eta_1 \| P \| = \lim_{n\to\infty} \eta_2 \| P \| = 0. \tag{39}$$

Dans le théorème 4

$$r = |\lambda| \, \eta_1 (1 + \| K^{-1}PK \|), \quad \| K^{-1} \| \leqslant \frac{1 + \| \widetilde{K}^{-1}P \| + \| \widetilde{K}^{-1}PK \|}{1 - r}. \tag{40}$$

Indiquons les conditions garantissant la réalisation des relations (39), donc la convergence de la suite des solutions approchées vers la solution exacte. Soit donnée une suite d'opérateurs $P_n$ projetant X sur les sous-espaces $\widetilde{X}_n$. A chaque $P_n$ correspond une équation approchée (36).

Théorème 6. *Si*

1) X *est un espace complet ;*
2) $P_n \to I$ *sur* X, *c'est-à-dire* $\lim_{n\to\infty} P_n x = x \; (x \in X)$ :
3) *l'opérateur* $H$ *est compact ;*

*alors les constantes* $\eta_1^{(n)}$ *et* $\eta_2^{(n)}$ *des conditions* II *et* III *peuvent être choisies telles que* $\eta_1^{(n)} \to 0$ *et* $\eta_2^{(n)} \to 0$ *pour* $n \to \infty$.

Démonstration. Soit $B$ la boule unité de l'espace X. L'ensemble $H(B)$ est compact, donc $P_n \to I$ uniformément sur $H(B)$ d'après le théorème de Guelfand (IX.1.4, tome 1). En posant

$$\eta_1^{(n)} = \sup_{z \in H(B)} \| P_n z - z \| \quad (n = 1, 2, \ldots),$$

on aura

$$\eta_1^{(n)} \to 0.$$

D'autre part, pour tout $x \in X$

$$\| P_n Hx - Hx \| \leqslant \eta_1^{(n)} \| x \|.$$

Comme $P_n Hx \in \widetilde{X}_n$, il résulte de cette inégalité que $\eta_1^{(n)}$ vérifie la condition II.

Pour $\eta_2^{(n)}$ on peut prendre la quantité

$$\eta_2^{(n)} = \frac{\| P_n y - y \|}{\| y \|}$$

qui tend manifestement vers 0.

Corollaire 1. *Si dans les conditions du théorème* $\lambda$ *n'est pas valeur caractéristique de l'opérateur* $H$, *alors les conditions du théorème* 3 *sont remplies et par conséquent la suite* $\{\widetilde{x}_n^*\}$ *des solutions approchées converge vers la solution exacte.*

En effet, (39) est réalisée puisque $\sup_n \| P_n \| < \infty$ d'après la condition 2) du théorème (VII.1.2, tome 1).

Corollaire 2. *Les valeurs caractéristiques de l'opérateur $H$ sont limites des suites de valeurs caractéristiques des opérateurs $\widetilde{H}_n$.*

Appesantissons-nous sur le cas où $\widetilde{X}$ est un espace de dimension $n$ finie. Chaque $\widetilde{x} \in \widetilde{X}$ se représente d'une seule façon sous la forme

$$\widetilde{x} = c_1\omega_1 + c_2\omega_2 + \ldots + c_n\omega_n, \tag{41}$$

où les éléments $\omega_1, \omega_2, \ldots, \omega_n$ forment une base dans $\widetilde{X}$.

Soit $f_1, f_2, \ldots, f_n$ un système de fonctionnelles linéaires dans $\widetilde{X}$, complet dans $\widetilde{X}$, c'est-à-dire tel que

$$f_j(\widetilde{x}) = 0 \quad (j = 1, 2, \ldots, n)$$

entraîne $\widetilde{x} = 0$.

L'équation approchée (36) est équivalente au système

$$f_j(PK\widetilde{x}) = f_j(Py) \quad (j = 1, 2, \ldots, n).$$

Si l'on cherche la solution de l'équation (36) sous la forme (41), on obtient pour les coefficients $c_k$ le système d'équations linéaires algébriques

$$\sum_{k=1}^{n} \alpha_{jk}c_k - \lambda \sum_{k=1}^{n} a_{jk}c_k = b_j \quad (j = 1, 2, \ldots, n), \tag{42}$$

où

$$\alpha_{jk} = f_j(\omega_k), \quad a_{jk} = f_j(PH\omega_k), \quad b_j = f_j(Py)$$
$$(j, k = 1, 2, \ldots, n).$$

Si le système de fonctionnelles $f_1, f_2, \ldots, f_n$ est biorthogonal à la base $\omega_1, \omega_2, \ldots, \omega_n$, le système (42) se simplifie:

$$c_j - \lambda \sum_{k=1}^{n} a_{jk}c_k = b_j \quad (j = 1, 2, \ldots, n) \tag{43}$$

En particulier, si $X$ est un espace hilbertien et $P$ un projecteur, en supposant que le système $\omega_1, \omega_2, \ldots, \omega_n$ est orthonormal on obtient

$$a_{jk} = f_j(PH\omega_k) = (PH\omega_k, \omega_j) = (H\omega_k, P\omega_j) = (H\omega_k, \omega_j)$$
$$(j, k = 1, 2, \ldots, n),$$
$$b_j = f_j(Py) = (Py, \omega_j) = (y, P\omega_j) = (y, \omega_j) \quad (j = 1, 2, \ldots, n),$$

et le système (42) s'écrit sous la forme

$$c_j - \lambda \sum_{k=1}^{n} c_k (H\omega_k, \omega_j) = (y, \omega_j) \quad (j = 1, 2, \ldots, n). \tag{44}$$

Le système (42) sera appelé système de la *méthode de Galerkin* dans la forme abstraite *).

Voici un autre cas particulier de la théorie générale, donnant lieu à de plus grandes simplifications.

Si l'espace X est confondu avec $\tilde{X}$, les conditions II et III sont manifestement réalisées avec $\eta_1 = \eta_2 = 0$. La condition I se transforme en la condition de proximité des opérateurs $H$ et $\tilde{H}$ (qui agissent alors dans le même espace)

$$\| H - \tilde{H} \| \leqslant \eta.$$

On remarquera que ce cas ne présente pas un intérêt notable du point de vue de la théorie générale, car tous les théorèmes peuvent être démontrés par des moyens élémentaires (cf. V.4.6, tome 1).

1.9. Il arrive fréquemment que l'équation approchée de telle ou telle méthode est considérée non pas dans un sous-espace de X mais dans un autre espace $\bar{X}$ qui, généralement, est isomorphe à un sous-espace de X.

On admettra que dans un sous-espace *complet* $\tilde{X} \subset X$ est défini un opérateur linéaire continu $\varphi_0$ appliquant bijectivement $\tilde{X}$ sur l'espace *complet* $\bar{X}$. Ces conditions assurent la continuité de l'opérateur inverse $\varphi_0^{-1}$. On supposera qu'il existe un opérateur linéaire continu $\varphi$, prolongement de $\varphi_0$ à X tout entier; donc, $\varphi$ est un opérateur linéaire continu de X sur $\bar{X}$, confondu avec $\varphi_0$ sur $\tilde{X}$.

Si l'on considère le projecteur $P$ de X sur $\tilde{X}$, introduit dans 1.1, pour $\varphi$ on peut prendre

$$\varphi = \varphi_0 P. \tag{45}$$

En multipliant les deux membres de cette égalité par $\varphi_0^{-1}$, on obtient

$$P = \varphi_0^{-1}\varphi. \tag{46}$$

Les éléments des espaces $\tilde{X}$ et $\bar{X}$ se correspondant par une bijection, l'équation approchée (2) peut être transformée en une équation équivalente, mais dans l'espace $\bar{X}$. Pour cela il faut remplacer $\tilde{x}$ par $\varphi_0^{-1}\bar{x}$ dans l'équation (2), puis appliquer l'opérateur $\varphi_0$ aux deux membres de l'équation. On obtient

$$\bar{x} - \varphi_0 \tilde{H} \varphi_0^{-1} \bar{x} = \varphi_0 P y.$$

Désignons par $\bar{H}$ l'opérateur de $\bar{X}$ dans $\bar{X}$

$$\bar{H} = \varphi_0 \tilde{H} \varphi_0^{-1}, \tag{47}$$

---

*) Pour les équations non linéaires la méthode de Galerkin est développée dans les ouvrages: Krasnosselski et autres et dans Vaïnikko.

et, en tenant compte de (45), mettons l'équation approchée sous la forme

$$\overline{K}\overline{x} \equiv \overline{x} - \lambda \overline{H}\overline{x} = \varphi y \quad (\overline{K} = \varphi_0 \widetilde{K} \varphi_0^{-1}). \tag{48}$$

La forme (48) de l'équation approchée étant fréquente dans les applications, nous allons énoncer les principaux théorèmes de la théorie générale en termes d'opérateurs $\overline{H}$, $\varphi$ (et $\overline{K}$). Pour cela il faut partout remplacer $P$ par son expression (46). Pour l'opérateur $\widetilde{H}$ ($\widetilde{K}$) il vient de (47)

$$\widetilde{H} = \varphi_0^{-1} \overline{H} \varphi_0 \quad (\widetilde{K} = \varphi_0^{-1} \overline{K} \varphi_0). \tag{49}$$

Attardons-nous sur la condition I. Dans les nouveaux termes elle devient

$$\| \varphi_0^{-1} \overline{H} \varphi_0 \widetilde{x} - \varphi_0^{-1} \varphi H \widetilde{x} \| \leqslant \eta \| \widetilde{x} \| \quad (\widetilde{x} \in \widetilde{X}).$$

Cette inégalité sera réalisée sous la condition

I bis. Pour tout $\widetilde{x} \in \widetilde{X}$ on a

$$\| \overline{H} \varphi_0 \widetilde{x} - \varphi H \widetilde{x} \| \leqslant \overline{\eta} \| \widetilde{x} \|.$$

Pour $\eta$, dans la condition I, il convient de prendre

$$\eta = \overline{\eta} \| \varphi_0^{-1} \|. \tag{50}$$

Les conditions II et III subsistent, car elles ne portent pas sur les opérateurs $\widetilde{H}$ et $P$.

Donnons maintenant une nouvelle formulation du théorème 1.

Théorème 1 bis. *Si sont réalisées les conditions* I bis, II, *l'opérateur $K$ admet un inverse continu et*

$$\overline{q} = |\lambda| \, [\overline{\eta} \| \varphi_0^{-1} \| + \| I - \varphi_0^{-1} \varphi \| \, \eta_1] \, \| K^{-1} \| < 1, \tag{51}$$

*alors l'opérateur $\overline{K}$ possède aussi un inverse continu $\overline{K}^{-1}$ et de plus*

$$\| \overline{K}^{-1} \| \leqslant \frac{\overline{N}}{1 - \overline{q}}, \tag{52}$$

*où*

$$\overline{N} = \| K^{-1} \| \, \| \varphi_0 \| \, \| \varphi_0^{-1} \|. \tag{53}$$

En effet, $\overline{K} = \varphi_0 \widetilde{K} \varphi_0^{-1}$ et, d'après le théorème 1, l'opérateur $\widetilde{K}$ admet un inverse $\widetilde{K}^{-1}$ et est justiciable de la majoration (4).

Si $\overline{x}^*$ est la solution de l'équation (48), la solution approchée $\widetilde{x}^*$ de l'équation (1) sera $\widetilde{x}^* = \varphi_0^{-1} \overline{x}^*$. L'erreur de la solution approchée peut être majorée avec le théorème 2 *mutatis mutandis.*

Théorème 2 bis. *Si sont réalisées les conditions* I bis, II *et* III, *l'opérateur inverse continu $\overline{K}^{-1}$ existe et l'équation* (1) *admet une solu-*

*tion $x^*$, alors*

$$\| x^* - \varphi_0^{-1}\bar{x}^* \| \leqslant \bar{p} \, \| x^* \|, \tag{54}$$

*où*

$$\bar{p} = (1+\varepsilon)\,|\lambda|\,\bar{\eta}\,\| \varphi_0^{-1}\bar{K}^{-1} \| + \varepsilon\,(1 + \| \varphi_0^{-1}\bar{K}^{-1}\varphi K \|), \tag{55}$$

*et*

$$\varepsilon \leqslant \eta_1\,|\lambda| + \eta_2\,\| K \|.$$

*Pour $\bar{p}$ on peut prendre également*

$$\bar{p} = 2\,|\lambda|\,\bar{\eta}\,\| \varphi_0^{-1} \|\,\| \varphi_0^{-1}\bar{K}^{-1}\varphi_0 \| + \varepsilon\,(1 + \| \varphi_0^{-1}\bar{K}^{-1}\varphi K \|). \tag{56}$$

La possibilité pour $\bar{p}$ de prendre la valeur (56) découle immédiatement de (13). L'égalité (55) implique une petite justification. En conservant les notations de la démonstration du théorème 2, on a d'après la condition I bis

$$\| \tilde{x} - \tilde{x}_0 \| = \| \varphi_0^{-1}\bar{K}^{-1}\bar{K}\varphi_0\tilde{x} - \varphi_0^{-1}\bar{K}^{-1}\varphi_0\varphi_0^{-1}\varphi K\tilde{x} \| \leqslant$$
$$\leqslant \| \varphi_0^{-1}\bar{K}^{-1} \|\,\| \bar{K}\varphi_0\tilde{x} - \varphi K\tilde{x} \| \leqslant |\lambda|\,\bar{\eta}\,\| \varphi_0^{-1}\bar{K}^{-1} \|\,\| \tilde{x} \|.$$

En utilisant cette majoration au lieu de (17), on obtient (55).

Signalons que les résultats mentionnés dans les remarques suivant le théorème 2 restent valables *mutatis mutandis.*

Formulons enfin le théorème de convergence de la suite des solutions approchées.

THÉORÈME 3 bis. *Si*

1) *l'opérateur $K$ admet un inverse continu;*

2) *pour chaque $n = 1, 2, \ldots$ sont réalisées les conditions* I bis, II *et* III *et de plus*

$$\lim_{n\to\infty} \bar{\eta}\,\| \varphi_0^{-1} \| = \lim_{n\to\infty} \eta_1\,\| \varphi_0^{-1}\varphi \| = \lim_{n\to\infty} \eta_2\,\| \varphi_0^{-1}\varphi \| = 0, \tag{57}$$

*alors, pour $n$ assez grand, l'équation approchée* (48) *admet une solution et la suite des solutions approchées converge vers la solution exacte. En outre*

$$\| x^* - \tilde{x}_n^* \| \leqslant \bar{Q}\bar{\eta}\,\| \varphi_0^{-1} \| + \bar{Q}_1\eta_1\,\| \varphi_0^{-1}\varphi \| + \bar{Q}_2\eta_2\,\| \varphi_0^{-1}\varphi \|, \tag{58}$$

*où $\tilde{x}_n^* = \varphi_0^{-1}\bar{x}_n^*$ et $\bar{Q}, \bar{Q}_1, \bar{Q}_2$ sont des constantes.*

En effet, grâce à (50) et (46) on conclut à la réalisation de toutes les conditions du théorème 3, d'où le résultat annoncé.

Dans les nouveaux termes la condition (25) du théorème 4 devient

$$\bar{r} = |\lambda|\,[(1 + |\lambda|\,\eta_1)\,\bar{\eta}\,\| \varphi_0^{-1}\bar{K}^{-1} \| + \eta_1\,\| \varphi_0^{-1}\bar{K}^{-1}\varphi K \|] < 1, \tag{59}$$

et la condition (26):

$$\| K^{-1} \| \leqslant \frac{1 + \| \varphi_0^{-1}\bar{K}^{-1}\varphi \| + |\lambda|\,\bar{\eta}\,\| \varphi_0^{-1}\bar{K}^{-1} \| + \| \varphi_0^{-1}\bar{K}^{-1}\varphi K \|}{1 - \bar{r}}. \tag{60}$$

Pour le prouver il faut refaire les mêmes raisonnements que pour la démonstration du théorème 4, en utilisant la majoration (54) du théorème 2 bis avec $\bar{p}$ de (55).

Une substitution directe dans (25) et (26) donne pour $\bar{r}$ la valeur

$$\bar{r} = |\lambda| \, [(1 + |\lambda| \, \eta_1) \, \bar{\eta} \, \| \varphi_0^{-1} \| \, \| \varphi_0^{-1} \bar{K}^{-1} \varphi_0 \| + \\ + \eta_1 (1 + \| \varphi_0^{-1} \bar{K}^{-1} \varphi K \|)], \tag{61}$$

et, au lieu de (60), on obtient la majoration

$$\| K^{-1} \| \leqslant \frac{1 + \| \varphi_0^{-1} \bar{K}^{-1} \varphi \| + |\lambda| \, \bar{\eta} \, \| \varphi_0^{-1} \| \, \| \varphi_0^{-1} \bar{K}^{-1} \varphi_0 \| + \| \varphi_0^{-1} \bar{K}^{-1} \varphi K \|}{1 - \bar{r}}, \tag{62}$$

où $\bar{r}$ a (61) pour expression.

## § 2. Equations réductibles à des équations de seconde espèce

**2.1.** Nous allons considérer des équations dont le second membre, bien qu'il ne contienne pas explicitement l'opérateur identité, est un opérateur linéaire continu de l'espace donné dans un autre espace normé, et qui grâce à la forme simple de la partie principale de cet opérateur peuvent être ramenées aux équations de seconde espèce examinées au § 1.

Soient $\mathbf{X}$ et $\mathbf{Y}$ des espaces normés, $\tilde{\mathbf{X}} \subset \mathbf{X}$ et $\tilde{\mathbf{Y}} \subset \mathbf{Y}$ des sous-espaces complets, $\Phi$ un opérateur linéaire continu projetant $\mathbf{Y}$ sur $\tilde{\mathbf{Y}}$.

Considérons l'équation exacte

$$K_1 x \equiv Gx - \lambda T x = y_1 \tag{1}$$

et l'équation approchée correspondante

$$\tilde{K}_1 \tilde{x} \equiv G\tilde{x} - \lambda \tilde{T} \tilde{x} = \Phi y_1. \tag{2}$$

$G$ et $T$ (et $K_1$) sont des opérateurs linéaires continus de $\mathbf{X}$ dans $\mathbf{Y}$, $\tilde{T}$ (et $\tilde{K}$), de $\tilde{\mathbf{X}}$ dans $\tilde{\mathbf{Y}}$. On admettra de plus que

1) l'opérateur $G$ admet un inverse continu;

2) l'opérateur $G$ établit une bijection de $\tilde{\mathbf{X}}$ sur $\tilde{\mathbf{Y}}$, c'est-à-dire $G(\tilde{\mathbf{X}}) = \tilde{\mathbf{Y}}$, donc $G^{-1}(\tilde{\mathbf{Y}}) = \tilde{\mathbf{X}}$.

Introduisons des conditions semblables aux conditions I à III du § 1.

I ter. Pour tout $\tilde{x} \in \tilde{\mathbf{X}}$ on a

$$\| \Phi T \tilde{x} - \tilde{T} \tilde{x} \| \leqslant \mu \, \| \tilde{x} \|.$$

II ter. Pour tout $x \in \mathbf{X}$ il existe $\tilde{y} \in \tilde{\mathbf{Y}}$ tel que

$$\| Tx - \tilde{y} \| \leqslant \mu_1 \, \| x \|.$$

III ter. Il existe un élément $\tilde{y}_1 \in \tilde{Y}$ tel que

$$\| y_1 - \tilde{y}_1 \| \leqslant \mu_2 \| y_1 \|.$$

Prouvons que sous les conditions posées, l'étude des équations (1) et (2) se ramène à celle d'équations de seconde espèce.

En effet, en appliquant l'opérateur $G^{-1}$ aux deux membres des équations (1) et (2) on obtient

$$Kx \equiv G^{-1}K_1 x \equiv x - \lambda G^{-1}Tx = G^{-1}y_1, \quad (3)$$

$$\tilde{K}\tilde{x} \equiv G^{-1}\tilde{K}_1 \tilde{x} \equiv \tilde{x} - \lambda G^{-1}\tilde{T}\tilde{x} = G^{-1}\Phi y_1. \quad (4)$$

Les équations (3) et (4) sont de la même forme que les équations (1) et (2) du § 1, les opérateurs $H$, $\tilde{H}$, $P$ ($K$, $\tilde{K}$) étant $G^{-1}T$, $G^{-1}\tilde{T}$, $G^{-1}\Phi G$ ($G^{-1}K_1$, $G^{-1}\tilde{K}_1$) et $y$ l'élément $G^{-1}y_1$.

Vérifions que les conditions I, II, III sont satisfaites.

I. D'après I ter on a

$$\| PH\tilde{x} - \tilde{H}\tilde{x} \| = \| G^{-1}\Phi G G^{-1}T\tilde{x} - G^{-1}\tilde{T}\tilde{x} \| \leqslant \| G^{-1} \| \, \mu \, \| \tilde{x} \|.$$

II. Pour $x \in X$ prenons $\tilde{y} \in \tilde{Y}$ d'après la condition II ter. En posant $\tilde{x} = G^{-1}\tilde{y}$ on aura

$$\| Hx - \tilde{x} \| = \| G^{-1}Tx - G^{-1}\tilde{y} \| \leqslant \| G^{-1} \| \, \| \mu_1 \| \, \| x \|.$$

III. Soit $\tilde{y}_1$ défini d'après III ter. Posons $\tilde{y} = G^{-1}\tilde{y}_1$. On a

$$\| y - \tilde{y} \| = \| G^{-1}y_1 - G^{-1}\tilde{y}_1 \| \leqslant \| G^{-1} \| \, \mu_2 \, \| y_1 \| \leqslant \| G^{-1} \| \, \| G \| \, \mu_2 \, \| y \|.$$

Donc, les conditions I, II, III du § 1 sont satisfaites, les constantes $\eta$, $\eta_1$, $\eta_2$ étant liées à $\mu$, $\mu_1$, $\mu_2$, par:

$$\eta = \mu \, \| G^{-1} \|, \quad \eta_1 = \mu_1 \, \| G^{-1} \|, \quad \eta_2 = \mu_2 \, \| G^{-1} \| \, \| G \|. \quad (5)$$

Remarque. Au lieu des conditions II ter et III ter on aurait pu exiger l'existence de $\tilde{x} \in \tilde{X}$ et $\tilde{y} \in \tilde{X}$, tels que

$$\| G^{-1}Tx - \tilde{x} \| \leqslant \mu_1' \, \| x \|, \quad \| G^{-1}y_1 - \tilde{y} \| \leqslant \mu_2' \, \| G^{-1}y_1 \|. \quad (6)$$

Les conditions II et III sont alors remplies avec $\eta_1 = \mu_1'$ et $\eta_2 = \mu_2'$.

Les conditions I, II et III étant satisfaites pour les équations (3) et (4), on peut, à partir des théorèmes du paragraphe précédent, déduire les théorèmes correspondants pour les équations (3) et (4) ou pour les équations équivalentes (1) et (2). Nous allons formuler ces théorèmes en détail moyennant toutefois une hypothèse simplificatrice.

Plus exactement supposons que dans les espaces X et Y les normes sont compatibles en vertu de l'application $G$:

$$\| x \|_X = \| Gx \|_Y, \quad \| y \|_Y = \| G^{-1}y \|_X \quad (x \in X, y \in Y). \quad (7)$$

L'application $G$ est manifestement isométrique sous cette condition; $\| G \| = \| G^{-1} \| = 1$. Donc

$$\| P \| = \| \Phi \|, \quad \| K \| = \| K_1 \|, \quad \| \tilde{K} \| = \| \tilde{K}_1 \|,$$
$$\| H \| = \| T \|, \quad \| \tilde{H} \| = \| \tilde{T} \|. \tag{8}$$

2.2. Formulons maintenant les principaux théorèmes du § 1 pour les équations (1) et (2).

Théorème 1 ter. *Si les conditions* I ter *et* II ter *sont satisfaites, l'opérateur inverse continu* $K_1^{-1}$ *existe et*

$$q = |\lambda| [\mu + \| I - \Phi \| \mu_1] \| K_1^{-1} \| < 1, \tag{9}$$

*alors l'opérateur* $\tilde{K}_1$ *admet aussi un inverse continu tel que*

$$\| \tilde{K}_1^{-1} \| \leqslant \frac{\| K_1^{-1} \|}{1-q}. \tag{10}$$

Théorème 2 ter. *Si les conditions* I ter, II ter *et* III ter *sont réunies, l'opérateur inverse continu* $\tilde{K}_1^{-1}$ *existe (ceci aura lieu en particulier, si les conditions du théorème* I ter *sont satisfaites) et l'équation* (1) *admet une solution* $x^*$, *alors*

$$\| x^* - \tilde{x}^* \| \leqslant p \| x^* \|, \tag{11}$$

*où* $\tilde{x}^*$ *est une solution de l'équation* (2) *et*

$$p = 2 |\lambda| \mu \| \tilde{K}_1^{-1} \| + (\mu_1 |\lambda| + \mu_2 \| K_1 \|) (1 + \| \tilde{K}_1^{-1} \Phi K_1 \|). \tag{12}$$

Remarque. Si existe un $\tilde{x} \in \tilde{X}$ tel que

$$\| x^* - \tilde{x} \| \leqslant \varepsilon \| x^* \|, \tag{13}$$

la majoration (11) est valable sans les conditions II ter et III ter, et pour

$$p = 2 |\lambda| \mu \| \tilde{K}_1^{-1} \| + \varepsilon (1 + \| \tilde{K}_1^{-1} \Phi K_1 \|). \tag{14}$$

Théorème 3 ter. *Si pour tout* $n = 1, 2, \ldots$ *les conditions* I ter, II ter *et* III ter *sont vérifiées, l'opérateur* $K_1$ *admet un inverse continu et*

$$\lim_{n\to\infty} \mu = \lim_{n\to\infty} \mu_1 \| \Phi \| = \lim_{n\to\infty} \mu_2 \| \Phi \| = 0, \tag{15}$$

*alors la suite des solutions approchées (c'est-à-dire des solutions de l'équation* (2)) *converge vers la solution* $x^*$ *de l'équation* (1). *De plus*

$$\| x^* - \tilde{x}_n^* \| \leqslant Q\mu + Q_1 \mu_1 \| \Phi \| + Q_2 \mu_2 \| \Phi \|. \tag{16}$$

Théorème 4 ter. *Si l'opérateur inverse continu* $\tilde{K}_1^{-1}$ *existe, les conditions* I ter *et* II ter *sont remplies et*

$$r = |\lambda| \mu (1 + |\lambda| \mu_1) \| \tilde{K}_1^{-1} \| + |\lambda| \mu_1 (1 + \| \tilde{K}_1^{-1} \Phi K_1 \|) < 1, \tag{17}$$

*l'opérateur $K_1$ admet un inverse à gauche continu et de plus*

$$\| K_1^{-1} \| \leqslant \frac{1 + \| \tilde{K}_1^{-1} \Phi \| + |\lambda| \, \mu \, \| \tilde{K}_1^{-1} \| + \| \tilde{K}_1^{-1} \Phi K_1 \|}{1 - r}. \tag{18}$$

**2.3.** Signalons enfin un cas particulier où les théorèmes se simplifient. Plus exactement, supposons que l'équation (2) est la projection de l'équation exacte. Autrement dit, on obtient l'équation approchée par application de l'opérateur $\Phi$ aux deux membres de l'équation (1):

$$\tilde{K}_1 \tilde{x} \equiv \Phi K_1 \tilde{x} \equiv G\tilde{x} - \lambda \Phi T \tilde{x} = \Phi y_1, \tag{19}$$

c'est-à-dire dans ce cas $\tilde{T} = \Phi T$.

Il est immédiat de voir que la condition I ter est alors vérifiée pour $\mu = 0$. En effet,

$$\| \tilde{T}\tilde{x} - \Phi T \tilde{x} \| = \| \Phi T \tilde{x} - \Phi T \tilde{x} \| = 0.$$

Donc, dans tous les énoncés des théorèmes on peut prendre $\mu = 0$.

Ajoutons encore un théorème déduit du théorème 1.6 et se rapportant exclusivement aux équations du type indiqué.

Théorème 6 ter. *Soient données une suite d'équations approchées de la forme* (19) *et les applications* $\Phi_n$,

1) *l'espace* Y *étant complet,*
2) $\Phi_n \to I$ *sur* Y,
3) *l'opérateur* $G^{-1}T$ *étant compact.*

*Si existe l'opérateur inverse* $K_1^{-1}$, *pour n assez grand les équations approchées admettent des solutions qui tendent vers la solution exacte.*

## § 3. Application aux systèmes infinis d'équations *)

**3.1.** Soit donné un système infini d'équations

$$\xi_j - \lambda \sum_{k=1}^{\infty} a_{jk} \xi_k = b_j \quad (j = 1, 2, \ldots), \tag{1}$$

tel que

$$\sum_{j,\,k=1}^{\infty} |a_{jk}|^2 < \infty, \quad \sum_{j=1}^{\infty} |b_j|^2 < \infty. \tag{2}$$

On demande une solution vérifiant la condition

$$\sum_{k=1}^{\infty} |\xi_k|^2 < \infty. \tag{3}$$

*) Les systèmes infinis sont accessibles dans Riesz [4], Hellinger et Toeplitz [1], Kantorovitch et Krylov.

L'une des méthodes de résolution du système (1) les plus répandues est la *méthode* dite *de réduction* qui consiste à remplacer le système infini (1) par un système de $n$ équations à $n$ inconnues

$$\xi_j-\lambda\sum_{k=1}^{n} a_{jk}\xi_k=b_j \quad (j=1, 2, \ldots, n). \tag{4}$$

La solution de ce système est considérée comme une solution approchée du système initial (1).

Nous nous intéressons à l'estimation de l'erreur de la solution approchée et à la convergence des solutions approchées vers la solution exacte pour $n \to \infty$.

Pour pouvoir appliquer la théorie générale on conviendra que $\mathbf{X} = \mathbf{l}^2$. Le système (1) s'écrit alors sous forme d'une seule équation dans $\mathbf{X}$:

$$Kx \equiv x - \lambda Hx = y \quad (x=\{\xi_n\},\ y=\{b_n\}). \tag{5}$$

$H$ représente ici un opérateur linéaire continu (et compact d'après XI.2.2, tome 1) dans $\mathbf{l}^2$, défini par la matrice du système:

$$z=Hx, \quad \zeta_j=\sum_{k=1}^{\infty} a_{jk}\xi_k \quad (j=1, 2, \ldots;\ x=\{\xi_n\},\ z=\{\zeta_n\}).$$

Pour $\overline{\mathbf{X}}$ on prend l'espace euclidien de dimension finie $\mathbf{l}_n^2$. Pour $\tilde{\mathbf{X}}$ il est naturel de prendre l'ensemble des éléments de $\mathbf{l}^2$, dont toutes les coordonnées sont nulles à partir de la $(n+1)$-ième. La signification des opérateurs $\varphi_0$ et $\varphi_0^{-1}$ est évidente, l'opérateur $\varphi$ associe à l'élément $x=\{\xi_m\}\in \mathbf{l}^2$ l'élément

$$\overline{x}=\varphi x=(\xi_1, \xi_2, \ldots, \xi_n)\in \mathbf{l}_n^2.$$

Il est évident que

$$\|\varphi\| = \|\varphi_0\| = \|\varphi_0^{-1}\| = 1.$$

Le système (4) peut maintenant s'écrire sous forme d'une seule équation dans $\overline{\mathbf{X}}$, soit

$$\overline{K}\overline{x} \equiv \overline{x}-\lambda\overline{H}\overline{x}=\varphi y, \tag{6}$$

où $H$ est défini par la matrice « tronquée »

$$A_n=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Vérifions que les conditions I bis, II et III sont réalisées. Pour un $\tilde{x} = (\xi_1, \xi_2, \ldots, \xi_n, 0, \ldots)$ quelconque de $\tilde{X}$ on a

$$\bar{z} = \varphi H\tilde{x} - \bar{H}\varphi_0\tilde{x} \quad (\bar{z} = (\zeta_1, \zeta_2, \ldots, \zeta_n)),$$

$$\zeta_j = \sum_{k=1}^{n} a_{jk}\xi_k - \sum_{k=1}^{n} a_{jk}\xi_k = 0 \quad (j = 1, 2, \ldots, n).$$

Donc, la condition I bis est satisfaite pour $\bar{\eta} = 0$.

Pour vérifier la condition II, prenons un $x = \{\xi_m\} \in l^2$ et posons $\tilde{x} = [Hx]_n = (\zeta_1, \zeta_2, \ldots, \zeta_n, 0, \ldots)$ *). On a

$$\|Hx - \tilde{x}\| = [\sum_{j=n+1}^{\infty} \sum_{k=1}^{\infty} |a_{jk}\xi_k|^2]^{1/2} \leqslant [\sum_{j=n+1}^{\infty} \sum_{k=1}^{\infty} |a_{jk}|^2 \sum_{k=1}^{\infty} |\xi_k|^2]^{1/2} = \\ = \eta_1 \|x\|,$$

où

$$\eta_1 = [\sum_{j=n+1}^{\infty} \sum_{k=1}^{\infty} |a_{jk}|^2]^{1/2}.$$

Il est clair d'après (2) que $\eta_1 \to 0$ pour $n \to \infty$. Enfin, en posant $\tilde{y} = [y]_n$, on obtient

$$\|y - \tilde{y}\| = [\sum_{j=n+1}^{\infty} |b_j|^2]^{1/2} = \left[\frac{\sum_{j=n+1}^{\infty} |b_j|^2}{\sum_{j=1}^{\infty} |b_j|^2}\right]^{1/2} \|y\|,$$

si bien que dans la condition III on peut prendre

$$\eta_2 = \left[\frac{\sum_{j=n+1}^{\infty} |b_j|^2}{\sum_{j=1}^{\infty} |b_j|^2}\right]^{1/2}$$

où $\eta_2 \to 0$ pour $n \to \infty$.

Ce cas est donc justiciable de la théorie exposée dans 1.9. En particulier, le théorème 3 bis nous permet de conclure que si $\lambda$ n'est pas valeur caractéristique du système (1) (i.e. n'est pas valeur caractéristique de l'opérateur $H$), alors pour les $n$ assez grands, le système (4) est soluble et les solutions approchées convergent vers

*) $[x]_n$ désigne l'élément tronqué, i.e. l'élément obtenu à partir de $x \in l^2$ en annulant ses coordonnées de rang $n+1$. Il est évident que $[x]_n = \varphi_0^{-1}\varphi x = Px$.

la solution exacte. La vitesse de convergence est définie par

$$\| x^* - \varphi_0^{-1}\bar{x}_n^* \| \leqslant Q_1 \left[ \sum_{j=n+1}^{\infty} \sum_{k=1}^{\infty} |a_{jk}|^2 \right]^{1/2} + Q_2 \left[ \frac{\sum\limits_{j=n+1}^{\infty} |b_j|^2}{\sum\limits_{j=1}^{\infty} |b_j|^2} \right]^{1/2},$$

où $x^* = (\xi_1^*, \xi_2^*, \ldots, \xi_n^*, \ldots)$ désigne la solution du système (1) et $\bar{x}_n^* = (\xi_1^{(n)}, \xi_2^{(n)}, \ldots, \xi_n^{(n)})$ celle du système (4).

D'où il est clair que $\xi_k^*$ diffère peu de $\xi_k^{(n)}$ pour $k = 1, 2, \ldots$ $\ldots, n$, et pour $k > n$ la coordonnée $\xi_k^*$ est petite. Il est également clair que

$$\lim_{n\to\infty} \xi_k^{(n)} = \xi_k^* \quad (k = 1, 2, \ldots).$$

Le théorème 1.4 (avec la condition (59) § 1) entraîne le

THÉORÈME 1. *Si le système* (4) *admet une solution unique et*

$$\bar{r} = |\lambda| \, \eta_1 \, \| \bar{K}^{-1} \| \, \| \varphi K \| < 1, \tag{7}$$

*alors* $\lambda$ *n'est pas valeur caractéristique du système* (1) *et*

$$\| K^{-1} \| \leqslant \frac{1 + \| \bar{K}^{-1} \| (1 + \| \varphi K \|)}{1 - \bar{r}}.$$

Ce théorème est intéressant en ce sens qu'il permet d'établir la résolubilité du système infini en fonction des données sur la solution du système fini ; de plus, toutes les quantités figurant dans l'inégalité (7) peuvent être trouvées sans difficultés, si bien que le critère de solubilité de ce théorème est très efficace.

La quantité $\eta_1$ a été définie plus haut. Pour $\| \varphi K \|$ on a la majoration

$$\| \varphi K \| \leqslant 1 + \| \varphi H \| \leqslant 1 + \left[ \sum_{j=1}^{n} \sum_{k=1}^{\infty} |a_{jk}|^2 \right]^{1/2}.$$

Enfin, si $A_n$ est une matrice symétrique, $\| \bar{K}^{-1} \|$ vaut

$$\| \bar{K}^{-1} \| = \max_{j=1, 2, \ldots, n} \frac{1}{|1 - \lambda \lambda_j^{(n)}|},$$

où $\lambda_j^{(n)}$ sont les valeurs propres de la matrice $A_n$. Dans le cas général, la détermination de $\| \bar{K}^{-1} \|$ est plus difficile (cf. V.2.8, tome 1).

REMARQUE 1. Au lieu des systèmes vérifiant la condition (2) on aurait pu considérer une classe plus vaste de systèmes pour lesquels l'opérateur $H$ est compact de $\mathbf{l}^2$ dans $\mathbf{l}^2$. On aurait alors déterminé $\eta_1$ par la formule

$$\eta_1 = \| H - \varphi_0^{-1} \varphi H \varphi_0^{-1} \varphi \|,$$

et comme toujours $\eta_1 \to 0$ pour $n \to \infty$ (XI.2.2, tome 1), si bien que ce cas est également justiciable de tous les théorèmes de la théorie générale.

REMARQUE 2. La théorie générale peut être appliquée à d'autres classes de systèmes infinis qui peuvent être considérés comme des équations fonctionnelles dans d'autres espaces de suites. Exemples: les systèmes de Riesz vérifiant la condition

$$\sum_{j=1}^{\infty} [\sum_{k=1}^{\infty} |a_{jk}|^{p/(p-1)}]^{p-1} < \infty \quad (1 < p < \infty),$$

$H$ étant un opérateur dans $\mathbf{l}^p$ ; les systèmes réguliers et complètement réguliers (dans l'espace $\mathbf{l}^\infty$).

## § 4. Application aux équations intégrales *)

4.1. Une des méthodes les plus efficaces de résolution numérique des équations intégrales consiste à remplacer l'équation intégrale par un système algébrique d'équations linéaires moyennant une formule de quadrature.

Soit donnée l'équation

$$x(s) - \lambda \int_0^1 h(s, t)\, x(t)\, dt = y(s). \tag{1}$$

En remplaçant l'intégrale par une formule de sommation numérique, construite sur les nœuds $t_1, t_2, \ldots, t_n$, soit

$$\int_0^1 x(t)\, dt = \sum_{k=1}^{n} A_k x(t_k), \tag{2}$$

et en exigeant que l'égalité (1) ait lieu uniquement en ces nœuds, on arrive au système

$$x(t_j) - \lambda \sum_{k=1}^{n} A_k h(t_j, t_k)\, x(t_k) = y(t_j) \quad (j = 1, 2, \ldots, n), \tag{3}$$

dont la solution définit une valeur approchée de la solution cherchée aux points $t_1, t_2, \ldots, t_n$.

Appliquons la théorie générale à l'estimation de l'erreur de cette méthode. Pour fixer les idées, on admettra que dans l'équation (1) le second membre est une fonction périodique continue (de période 1) et que le noyau $h(s, t)$ est une fonction périodique continue (de période 1 aussi) en $t$ et en $s$. Sous ces conditions, la solution sera

*) Pour la résolution approchée des équations intégrales et la bibliographie relative à ce sujet consulter Kantorovitch et Krylov.

7—01292

également une fonction continue périodique. L'équation (1) sera traitée comme une équation fonctionnelle dans l'espace $X = \tilde{C}$ des fonctions périodiques continues. Le système algébrique (3) sera considéré comme une équation fonctionnelle approchée dans l'espace $\tilde{X} = l_n^\infty$,

$$\bar{K}\bar{x} \equiv \bar{x} - \lambda \bar{H}\bar{x} = \varphi y. \tag{4}$$

Les nœuds et les coefficients de la formule de quadrature (2) sont choisis de la façon suivante:

$$A_k = \frac{1}{n}, \quad t_k = \frac{2k-1}{2n} \quad (k = 1, 2, \ldots, n),$$

c'est-à-dire nous considérons la formule des rectangles moyens, qui est l'une des plus commodes dans le cas de fonctions périodiques. Ceci étant, l'opérateur $\bar{H}$ de (4) est déterminé par la matrice

$$\left\| \begin{array}{cccc} \frac{1}{n} h\left(\frac{1}{2n}, \frac{1}{2n}\right) & \frac{1}{n} h\left(\frac{1}{2n}, \frac{3}{2n}\right) & \cdots & \frac{1}{n} h\left(\frac{1}{2n}, \frac{2n-1}{2n}\right) \\ \frac{1}{n} h\left(\frac{3}{2n}, \frac{1}{2n}\right) & \frac{1}{n} h\left(\frac{3}{2n}, \frac{3}{2n}\right) & \cdots & \frac{1}{n} h\left(\frac{3}{2n}, \frac{2n-1}{2n}\right) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{n} h\left(\frac{2n-1}{2n}, \frac{1}{2n}\right) & \frac{1}{n} h\left(\frac{2n-1}{2n}, \frac{3}{2n}\right) & \cdots & \frac{1}{n} h\left(\frac{2n-1}{2n}, \frac{2n-1}{2n}\right) \end{array} \right\|. \tag{5}$$

L'opérateur $\varphi$ est défini comme suit:

$$\varphi x = (x(t_1), x(t_2), \ldots, x(t_n)), \quad \|\varphi\| = 1 \quad (x \in \tilde{C}).$$

S'agissant de l'espace $\tilde{X}$ et de l'application $\varphi_0^{-1}$ qui lui est rattachée, nous les définissons de deux manières distinctes, ce qui nous conduit à deux estimations de l'erreur.

Prenons d'abord pour $\tilde{X}$ l'ensemble des fonctions périodiques continues, linéaires dans chaque intervalle de la forme $[t_k, t_{k+1}]$, où $t_k = \frac{2k-1}{2n}$ $(k = 0, \pm 1, \pm 2, \ldots)$. Une telle fonction est définie par ses valeurs aux points $t_1, t_2, \ldots, t_n$, ce qui définit l'application $\varphi_0^{-1}$:

$$\tilde{x} = \varphi_0^{-1}\bar{x},$$

$$(\bar{x} = (\xi_1, \ldots, \xi_n) \in l_n^\infty; \quad \tilde{x}(t_k) = \xi_k; \quad k = 1, 2, \ldots, n; \quad \tilde{x} \in \tilde{X}).$$

Et de plus $\|\varphi_0^{-1}\| = 1$.

Notons $\omega_s(\delta)$ le module de continuité de la fonction $h(s, t)$ en tant que fonction de $s$:

$$\omega_s(\delta) = \sup |h(s + \sigma, t) - h(s, t)| \quad (0 \leqslant s, t \leqslant 1, |\sigma| \leqslant \delta);$$

définissons $\omega_t(\delta)$ de façon analogue.

Pour vérifier la condition I bis, estimons l'erreur de la formule de quadrature (2) pour une fonction de la forme $z(t)\,\tilde{x}(t)$, où $z$ est une fonction périodique dont le module de continuité est $\leqslant \omega(\delta)$, et $\tilde{x} \in \tilde{X}$. Compte tenu de la périodicité, on a

$$\left|\int_0^1 z(t)\,\tilde{x}(t)\,dt - \sum_{k=1}^{n} \frac{1}{n} z(t_k)\,\tilde{x}(t_k)\right| =$$

$$= \left|\int_{t_0}^{t_n} z(t)\,\tilde{x}(t)\,dt - \sum_{k=0}^{n-1} \frac{1}{2n}\left(z(t_k)\,\tilde{x}(t_k) + z(t_{k+1})\,\tilde{x}(t_{k+1})\right)\right| =$$

$$= \left|\sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \left[z(t) - z\left(\frac{t_k + t_{k+1}}{2}\right)\right]\tilde{x}(t)\,dt + \right.$$

$$+ \sum_{k=0}^{n} \frac{1}{2n}\left[\tilde{x}(t_k)\left(z\left(\frac{t_k+t_{k+1}}{2}\right) - z(t_k)\right) + \right.$$

$$\left. \left. + \tilde{x}(t_{k+1})\left(z\left(\frac{t_k+t_{k+1}}{2}\right) - z(t_{k+1})\right)\right]\right| \leqslant 2\omega\left(\frac{1}{2n}\right)\|\tilde{x}\|,$$

puisque

$$\int_{t_k}^{t_{k+1}} \tilde{x}(t)\,dt = \frac{1}{2n}\left(\tilde{x}(t_k) + \tilde{x}(t_{k+1})\right) \quad (k = 0, 1, \ldots, n-1)$$

en raison de la linéarité de la fonction $\tilde{x}(t)$ dans chaque intervalle. En posant $z(t) = h(t_j, t)$ dans l'inégalité précédente, on obtient

$$\|\varphi H\tilde{x} - \overline{H}\varphi_0\tilde{x}\| = \max_{j=1,2,\ldots,n}\left|\int_0^1 h(t_j, t)\,\tilde{x}(t)\,dt - \right.$$

$$\left. - \frac{1}{n}\sum_{k=1}^{n} h(t_j, t_k)\,\tilde{x}(t_k)\right| \leqslant 2\omega_t\left(\frac{1}{2n}\right)\|\tilde{x}\|.$$

Donc, la condition I bis est vérifiée pour

$$\overline{\eta} = 2\omega_t\left(\frac{1}{2n}\right).$$

Passons à la condition II. Nous prouvons tout d'abord que, si $z \in \tilde{C}$ est une fonction quelconque dont le module de continuité est $\leqslant \omega(\delta)$, alors

$$\|z - \tilde{z}\| \leqslant \omega\left(\frac{1}{n}\right),$$

où $\tilde{z} = \varphi_0^{-1}\varphi z$, c'est-à-dire $\tilde{z}$ est une fonction linéaire par morceaux prenant les mêmes valeurs que $z$ aux points $t_1, t_2, \ldots, t_n$. En effet, si $t_j \leqslant s \leqslant t_{j+1}$, alors

$$|z(s)-\tilde{z}(s)| = |z(s)-[(t_{j+1}-s)z(t_j)+(s-t_j)z(t_{j+1})]n| \leqslant$$
$$\leqslant n[|(t_{j+1}-s)(z(s)-z(t_j))|+|(s-t_j)(z(s)-z(t_{i+1}))|] \leqslant$$
$$\leqslant n\omega\left(\frac{1}{n}\right)(t_{j+1}-t_j) = \omega\left(\frac{1}{n}\right),$$

d'où ce que nous voulions, puisque $s$ est arbitraire. Estimons maintenant le module de continuité de $z = Hx$ $(x \in \tilde{C})$:

$$|z(s)-z(s')| \leqslant \int_0^1 |h(s,t)-h(s',t)|\,|x(t)|\,dt \leqslant \omega_s(\delta)\,\|x\|$$
$$(|s-s'| \leqslant \delta).$$

Donc, pour $\omega(\delta)$ on peut prendre

$$\omega(\delta) = \omega_s(\delta)\,\|x\|$$

et, d'après ce qui précède, si l'on désigne $\tilde{z} = \varphi_0^{-1}\varphi Hx$, on a

$$\|Hx-\tilde{z}\| \leqslant \omega_s\left(\frac{1}{n}\right)\|x\|.$$

Donc, la condition II est satisfaite avec

$$\eta_1 = \omega_s\left(\frac{1}{n}\right).$$

Enfin, en appliquant ce qui vient d'être dit à l'élément $y$, on trouve un $\tilde{y} \in \tilde{X}$, tel que

$$\|y-\tilde{y}\| \leqslant \bar{\omega}\left(\frac{1}{n}\right),$$

où $\bar{\omega}(\delta)$ est le module de continuité de la fonction $y$. D'où il est clair que la condition III est satisfaite avec

$$\eta_2 = \frac{1}{\|y\|}\,\bar{\omega}\left(\frac{1}{n}\right).$$

Le noyau $h(s, t)$ et le second membre $y(s)$ de l'équation (1) étant continus, les quantités $\bar{\eta}$, $\eta_1$, $\eta_2$ tendent vers 0 pour $n \to \infty$ et $\|\varphi\| = \|\varphi_0^{-1}\| = 1$, donc on peut se servir des théorèmes de la théorie générale. En particulier, si $\lambda$ n'est pas valeur caractéristique de l'équation (1), le théorème 1.3 bis nous dit que le système (3) est soluble (pour les $n$ assez grands) et que les solutions approchées $\varphi_0^{-1}\bar{x}_n^*$ (fonctions linéaires par morceaux construites avec les valeurs de $\{\xi_k\}$ tirées du système (3)) convergent vers la solution exacte.

Il importe de signaler pour la suite que d'après le théorème 1.1 bis, la norme de l'opérateur inverse $\| \bar{K}^{-1} \|$ est bornée indépendamment de $n$.

Si $h(s, t)$ et $y(s)$ vérifient une condition de Lipschitz dans le rapport $\alpha$, d'après le théorème 1.3 bis la vitesse de convergence des solutions approchées peut être estimée comme suit:

$$\| x^* - \varphi_0^{-1} \bar{x}_n^* \| = O\left(\frac{1}{n^\alpha}\right).$$

En effet, dans ce cas $\omega_s(\delta)$, $\omega_t(\delta)$ et $\bar{\omega}(\delta)$ sont de la forme $O(\delta^\alpha)$.

4.2. Considérons l'estimation de l'erreur de la même méthode pour un autre sous-espace $\tilde{X}$.

Pour simplifier nous admettrons que $n$ est impair: $n = 2m + 1$. Pour $\tilde{X}$ nous prenons l'ensemble des polynômes trigonométriques d'ordre $\leqslant m$, c'est-à-dire $\tilde{x} \in \tilde{X}$ signifie que

$$\tilde{x}(t) = \frac{a_0}{2} + \sum_{k=1}^{m} (a_k \cos 2k\pi t + b_k \sin 2k\pi t).$$

Il est clair que l'opérateur $\varphi_0^{-1}$ associe à $\bar{x} = (\xi_1, \xi_2, \ldots, \xi_n) \in l_n^\infty$ le polynôme d'interpolation prenant les valeurs $\xi_1, \xi_2, \ldots, \xi_n$ aux points $t_1, t_2, \ldots, t_n$. D'après un lemme de Bernstein *), si un polynôme trigonométrique $\tilde{x}$ admet la majoration

$$|\tilde{x}(t_j)| \leqslant 1 \quad (j = 1, 2, \ldots, n)$$

aux points $t_1, t_2, \ldots, t_n$, alors

$$|\tilde{x}(t)| \leqslant A \ln m + B \quad (0 \leqslant t \leqslant 1).$$

D'où

$$\| \varphi_0^{-1} \| \leqslant A \ln m + B. \tag{6}$$

Désignons par $E_m^t$ et $E_m^s$ les meilleures approximations de $h(s, t)$ comme fonction de $t$ et de $s$ respectivement par des polynômes trigonométriques d'ordre $m$ en $t$ et $s$, autrement dit nous admettons l'existence de fonctions

$$h_1(s, t) = \frac{a_0'(s)}{2} + \sum_{k=1}^{m} [a_k'(s) \cos 2k\pi t + b_k'(s) \sin 2k\pi t],$$

$$h_2(s, t) = \frac{a_0''(t)}{2} + \sum_{k=1}^{m} [a_k''(t) \cos 2k\pi s + b_k''(t) \sin 2k\pi s],$$

*) Cf. Natanson [I], page 542.

telles que

$$|h(s,\ t)-h_1(s,\ t)|\leqslant E_m^t, \quad |h(s,\ t)-h_2(s,\ t)|\leqslant E_m^s \quad (0\leqslant s,\ t\leqslant 1).$$

Pour la vérification de la condition I bis, on remarquera tout d'abord que la formule de quadrature (2) est exacte si la fonction à intégrer est un polynôme trigonométrique d'ordre $n-1=2m$ *).

$h_1(s,\ t)\ \tilde{x}\ (t)$ étant un tel polynôme en $t$, on a

$$\int_0^1 h_1(s,\ t)\ \tilde{x}\ (t)\ dt = \frac{1}{n}\sum_{k=1}^{n} h_1(s,\ t_k)\ \tilde{x}\ (t_k) \quad (0\leqslant s\leqslant 1).$$

Donc

$$\|\varphi H\tilde{x}-\overline{H}\varphi_0\tilde{x}\| =$$

$$= \max_{j=1,\,2,\,\ldots,\,n}\left|\int_0^1 h(t_j,\ t)\ \tilde{x}\ (t)\ dt-\frac{1}{n}\sum_{k=1}^{n} h(t_j,\ t_k)\tilde{x}(t_k)\right|\leqslant$$

$$\leqslant \max_{j=1,\,2,\,\ldots,\,n}\left|\int_0^1 [h(t_j,\ t)-h_1(t_j,\ t)]\ \tilde{x}\ (t)\ dt\right|+$$

$$+\max_{j=1,\,2,\,\ldots,\,n}\frac{1}{n}\left|\sum_{k=1}^{n}[h(t_j,\ t_k)-h_1(t_j,\ t_k)]\ \tilde{x}\ (t_k)\right|\leqslant 2E_m^t\,\|\tilde{x}\|.$$

Donc, la condition I bis est satisfaite pour

$$\overline{\eta}=2E_m^t.$$

Construisons la fonction **)

$$\tilde{x}(s)=\int_0^1 h_2(s,\ t)\,x(t)\,dt$$

pour vérifier la condition II afin d'approximer l'élément $z=Hx$. Il est clair, de la forme de $h_2(s,\ t)$, que la fonction $\tilde{x}(s)$ est un polynôme trigonométrique d'ordre $m$, c'est-à-dire $\tilde{x}\in\tilde{X}$. D'autre part

$$\|Hx-\tilde{x}\|=\max_s\left|\int_0^1 [h(s,\ t)-h_2(s,\ t)]\ x(t)\,dt\right|\leqslant E_m^s\,\|x\|.$$

---

*) Cf. Natanson [I], page 611.

**) On montre sans peine que la fonction $h_2(s,\ t)$ peut toujours être choisie continue en $t$, si bien que l'intégrale existe. Du reste, cette restriction n'est pas indispensable si l'intégrale est comprise au sens de Banach (cf. II.4.2, tome 1), car elle est définie pour toute fonction bornée.

Donc, la condition II est réalisée pour

$$\eta_1 = E_m^s.$$

Enfin, la condition III est satisfaite si l'on pose

$$\eta_2 = \frac{1}{\| y \|} E_m(y),$$

où $E_m(y)$ est la meilleure approximation de la fonction $y$ par des polynômes trigonométriques d'ordre $m$.

Grâce au théorème 1.2 bis, ces majorations nous permettent d'évaluer l'écart entre la solution approchée et la solution exacte:

$$\| x^* - \varphi_0^{-1} \bar{x}^* \| \leqslant \bar{p} \| x^* \|,$$

$$\bar{p} = 4 |\lambda| E_m^t \| \varphi_0^{-1} \| \| \bar{K}^{-1} \| +$$
$$+ \left( |\lambda| E_m^s + \frac{E_m(y) \| K \|}{\| y \|} \right) (1 + \| \varphi_0^{-1} \| \| \bar{K}^{-1} \| \| K \|).$$

On remarquera que $\| \bar{K}^{-1} \|$ est bornée indépendamment de $n$ (cf. 4.1).

Supposons que $h(s, t)$ admet des dérivées jusqu'à l'ordre $\nu$ compris par rapport à $s$ et $t$, les dérivées $\frac{\partial^\nu h}{\partial s^\nu}$ et $\frac{\partial^\nu h}{\partial t^\nu}$ vérifiant une condition de Lipschitz dans le rapport $\alpha$; la fonction $y(s)$ est assujettie aux mêmes conditions. Alors, d'après un théorème classique de Jackson (cf. Natanson [I], page 121)

$$E_m^t = O\left(\frac{1}{m^{\nu+\alpha}}\right) = O\left(\frac{1}{n^{\nu+\alpha}}\right), \quad E_m^s = O\left(\frac{1}{m^{\nu+\alpha}}\right) = O\left(\frac{1}{n^{\nu+\alpha}}\right),$$
$$E_m(y) = O\left(\frac{1}{m^{\nu+\alpha}}\right) = O\left(\frac{1}{n^{\nu+\alpha}}\right),$$

et comme d'après (6)

$$\| \varphi_0^{-1} \| = O(\ln m) = O(\ln n),$$

le théorème 1.3 bis nous dit que la vitesse de décroissance de l'erreur de la solution approchée est majorée par

$$\| x^* - \varphi_0^{-1} \bar{x}_n^* \| = O\left(\frac{\ln n}{n^{\nu+\alpha}}\right).$$

En utilisant le théorème 1.4, on peut, grâce aux résultats de la résolution approchée, déterminer le domaine possible de disposition des valeurs caractéristiques de l'équation (1).

Signalons enfin qu'on aurait pu par des raisonnements analogues évaluer l'erreur de la même méthode pour d'autres formules de quadrature. En particulier, pour le cas non périodique, en se servant de la formule de Gauss il aurait fallu faire appel à une interpolation par des polynômes algébriques et aux théorèmes d'approximation de fonctions par des polynômes algébriques.

**4.3.** Voyons d'autres méthodes de résolution approchée de l'équation intégrale (1), et, tout d'abord, la *méthode des moments*. Cette méthode consiste à chercher la solution approchée de l'équation (1) sous la forme

$$\tilde{x}(t)=\sum_{k=1}^{n} c_k\omega_k(t) \quad (0\leqslant t\leqslant 1), \tag{7}$$

où $\omega_k$ $(k = 1, 2, \ldots)$ sont des fonctions d'un système orthonormal complet. Les coefficients $c_k$ sont déterminés à partir du système

$$c_j-\lambda\sum_{k=1}^{n} c_k\int_0^1\int_0^1 h(s, t)\,\omega_j(s)\,\omega_k(t)\,ds\,dt=\int_0^1 y(s)\,\omega_j(s)\,ds$$

$$(j=1, 2, \ldots, n), \tag{8}$$

qui remplace la condition de nullité de la différence $Kx - y$ par la condition d'orthogonalité de $Kx - y$ aux $n$ premières fonctions du système orthonormal donné.

Dans l'hypothèse que le noyau $h(s, t)$ de l'équation (1) est tel que

$$\int_0^1\int_0^1 \left|h(s, t)\right|^2 ds\,dt<\infty,$$

prenons pour $\mathbf{X}$ l'espace $\mathbf{L}^2(0, 1)$ et pour $\tilde{\mathbf{X}}$ le sous-espace des éléments de la forme (7). Définissons l'opérateur $P$ comme un projecteur sur $\tilde{\mathbf{X}}$, i.e. $\tilde{x} = Px$ signifie que

$$\tilde{x}=\sum_{k=1}^{n}(x, \omega_k)\,\omega_k.$$

Le système (8) peut être écrit sous forme d'une équation dans l'espace $\tilde{\mathbf{X}}$

$$\tilde{x}-PH\tilde{x}=Py.$$

Nous sommes donc placés dans les conditions mentionnées dans 1.8. De plus, le système $\{\omega_k\}$ étant complet, les conditions du théorème 1.6 sont réalisées et, par suite, les solutions approchées convergent vers la solution exacte. La vitesse de convergence est déterminée par les propriétés du noyau $h(s, t)$ et du second membre $y(s)$ ; elle dépend de la vitesse de convergence des développements orthogonaux de ces fonctions suivant les fonctions du système. Traitons le cas où le noyau est une fonction périodique vérifiant la condition de Lipschitz dans le rapport $\alpha+\frac{1}{2}$ $(\alpha>0)$ par rapport à $s$. Imposons les mêmes conditions à $y(s)$. Pour système orthonormal $\{\omega_k\}$ prenons un système de fonctions trigonométriques. Le théorème de Jackson

nous donne immédiatement les estimations

$$\eta_1 = O(n^{-1/2-\alpha}), \quad \eta_2 = O(n^{-1/2-\alpha}),$$

donc

$$\| x^* - \tilde{x}_n^* \|_{L^2} = O(\eta_1 + \eta_2) = O(n^{-1/2-\alpha}). \tag{9}$$

Prouvons que dans ce cas les solutions approchées convergent uniformément vers la solution exacte.

Soit $\tilde{x} = \sum_{k=1}^{n} c_j \omega_j$. Evaluons $\|\tilde{x}\|_C$ en fonction de $\|\tilde{x}\|_{L^2}$. Comme

$$\| \omega_j \|_C \leqslant M \quad (j = 1, 2, \ldots),$$

en appliquant l'inégalité de Cauchy-Bouniakovski on obtient

$$\| \tilde{x} \|_C \leqslant M \sum_{k=1}^{n} |c_j| \leqslant M \sqrt{n} \, [\sum_{j=1}^{n} |c_j|^2]^{1/2} = M \sqrt{n} \, \| x \|_{L^2}.$$

Appliquons cette majoration à la différence des deux approximations $\tilde{x}_n^*$ et $\tilde{x}_{2^k}^*$, où $2^{k-1} \leqslant n \leqslant 2^k$. On obtient

$$\| \tilde{x}_{2^k}^* - \tilde{x}_n^* \|_C \leqslant M 2^{k/2} \| \tilde{x}_{2^k}^* - \tilde{x}_n^* \|_{L^2}.$$

La majoration (9) entraîne

$$\| \tilde{x}_{2^k}^* - \tilde{x}_n^* \|_{L^2} \leqslant \| \tilde{x}_{2^k}^* - x^* \|_{L^2} + \| x^* - \tilde{x}_n^* \|_{L^2} \leqslant$$
$$\leqslant \frac{C}{2^{\frac{k}{2}+\alpha k}} + \frac{C}{n^{\frac{1}{2}+\alpha}} \leqslant \frac{C_2}{2^{\frac{k}{2}+\alpha k}},$$

donc

$$\| \tilde{x}_{2^k}^* - \tilde{x}_n^* \|_C \leqslant \frac{C_2}{2^{\alpha k}}. \tag{10}$$

En particulier la dernière majoration est valable pour $n = 2^{k-1}$, c'est-à-dire

$$\| \tilde{x}_{2^k}^* - \tilde{x}_{2^{k-1}}^* \|_C \leqslant \frac{C_2}{2^{\alpha k}}.$$

D'où la convergence uniforme de la série

$$\tilde{x}_1^*(t) + \sum_{k=1}^{\infty} [\tilde{x}_{2^k}^*(t) - \tilde{x}_{2^{k-1}}^*(t)] = \lim_{k \to \infty} \tilde{x}_{2^k}^*(t) = x^*(t).$$

Compte tenu de (10) on trouve que

$$\lim_{n \to \infty} \tilde{x}_n^*(t) = x^*(t)$$

uniformément. De plus, il est évident que

$$\| x^* - \tilde{x}_n^* \|_C = O(n^{-\alpha}).$$

A noter que si l'on traite l'équation intégrale (1) comme une équation dans l'espace **C** [0, 1], le théorème 1.3 permet d'obtenir un résultat plus puissant que le précédent. Plus exactement, si le noyau (par rapport à $s$) et le second membre $y(s)$ vérifient la condition de Lipschitz dans le rapport $\beta > 0$, les solutions approchées convergent uniformément vers la solution exacte et

$$\| x^* - \tilde{x}_n^* \|_C = O(n^{-\beta} \ln n).$$

4.4. Voyons maintenant la méthode de substitution au noyau d'un noyau proche (en particulier dégénéré). Comme son nom l'indique, cette méthode consiste à substituer à l'équation (1) l'équation

$$\tilde{x}(s) - \lambda \int_0^1 \tilde{h}(s, t) \tilde{x}(t)\, dt = y(s), \tag{11}$$

dont le noyau $\tilde{h}(s, t)$ est proche de $h(s, t)$ et dont la solution est connue.

En écrivant l'équation (11) comme une équation fonctionnelle dans le même espace que l'équation (1):

$$\tilde{K}\tilde{x} \equiv \tilde{x} - \lambda \tilde{H}\tilde{x} = y,$$

on retombe dans la situation décrite à la fin de 1.8. On a indiqué là-bas que la condition I se ramenait à l'inégalité

$$\| H - \tilde{H} \| \leqslant \eta, \tag{12}$$

et les conditions II et III étaient satisfaites pour $\eta_1 = \eta_2 = 0$.

Prenons pour espace $\mathbf{X} = \tilde{\mathbf{X}}$ l'espace $\mathbf{L}^\infty(0, 1)$ des fonctions mesurables bornées et mettons (12) sous la forme

$$\int_0^1 | h(s, t) - \tilde{h}(s, t) |\, dt \leqslant \eta$$

$$(0 \leqslant s \leqslant 1).$$

On évalue sans peine la norme de l'opérateur $\tilde{K}^{-1}$ si l'on connaît la résolvante $\tilde{\Gamma}(\lambda; s, t)$ de l'équation approchée (11). Comme l'opérateur inverse $\tilde{K}^{-1}$ est de la forme

$$z = \tilde{K}^{-1} y,$$

$$z(s) = y(s) + \lambda \int_0^1 \tilde{\Gamma}(\lambda; s, t)\, y(t)\, dt,$$

alors

$$\| \widetilde{K}^{-1} \| \leqslant 1 + |\lambda| B,$$

où $B$ est déterminé à partir de

$$\int_0^1 |\widetilde{\Gamma}(\lambda;\ s,\ t)|\, dt \leqslant B \qquad (0 \leqslant s \leqslant 1).$$

En se servant de la majoration (19) § 1, on obtient celle de l'écart entre la solution approchée et la solution exacte

$$|x^*(t) - \widetilde{x}^*(t)| \leqslant |\lambda| \eta (1 + |\lambda| B) \| x^* \|. \tag{13}$$

On obtient une majoration analogue en permutant le rôle des deux équations.

(13) entraîne la convergence uniforme des solutions approchées vers la solution exacte sous réserve de la convergence des noyaux.

La théorie générale peut être appliquée à une méthode de résolution consistant à chercher la solution sous la forme (7), l'égalité des deux membres de l'équation n'étant exigée qu'aux points donnés. Nous ne développerons pas cette méthode (cf. Kantorovitch [9]) car une méthode similaire est étudiée pour les équations différentielles dans le paragraphe suivant.

## § 5. Application aux équations différentielles ordinaires

5.1. Etudions tout d'abord la *méthode d'interpolation* appelée encore *méthode de collocation* *) dans le cas du problème aux limites pour l'équation

$$\frac{d^{2m}x}{dt^{2m}} - \lambda \left[ p_1 \frac{d^{2m-1}x}{dt^{2m-1}} + \ldots + p_{2m} x \right] = y, \tag{1}$$

avec les conditions

$$\begin{aligned} x(a) = x'(a) = \ldots = x^{(m-1)}(a) = 0, \\ x(b) = x'(b) = \ldots = x^{(m-1)}(b) = 0. \end{aligned} \tag{2}$$

On cherchera la solution approchée sous la forme

$$\widetilde{x}(t) = (t-a)^m (b-t)^m \sum_{k=1}^{n} c_k t^{k-1}, \tag{3}$$

si bien que les conditions aux limites pour $\widetilde{x}$ sont satisfaites; on détermine les coefficients $c_1,\ c_2,\ \ldots,\ c_n$ en réalisant (1) en un systè-

*) Il semble que la méthode de collocation soit due à Kantorovitch [1]. La démonstration de la convergence sur la base des théorèmes des §§ 1, 2 est effectuée dans l'article de Karpilovskaïa [1].

me de points (nœuds) donnés $t_1, t_2, \ldots, t_n$:

$$\left\{\frac{d^{2m}\tilde{x}}{dt^{2m}} - \lambda\left[p_1 \frac{d^{2m-1}\tilde{x}}{dt^{2m-1}} + \ldots + p_{2m}\tilde{x}\right]\right\}_{t=t_j} = y(t_j). \tag{4}$$

Pour pouvoir appliquer la théorie générale exposée au § 2, on traitera l'équation différentielle (1) comme une équation fonctionnelle dans l'espace $X = C^{(2m)}[a, b]$ des fonctions $2m$-continûment dérivables, vérifiant les conditions (2). La norme de l'espace X est définie plus bas.

Pour $\tilde{X}$, nous prenons l'ensemble des fonctions de la forme (3). Pour Y, l'espace $C[a, b]$ des fonctions continues sur l'intervalle $[a, b]$, la norme étant définie de la façon usuelle. Enfin, pour $\tilde{Y}$, l'ensemble des polynômes de degré $n - 1$.

Définissons l'application $\Phi$ de Y sur $\tilde{Y}$ en prenant pour $\tilde{y} = \Phi(y)$ le polynôme d'interpolation de degré $n - 1$ qui a aux points $t_1, t_2, \ldots, t_n$ les mêmes valeurs que $y$. On sait que dans le cas de nœuds de Tchébychev

$$\|\Phi\| \leqslant A \ln n + B. \tag{5}$$

On sait aussi que dans le cas de nœuds de Gauss

$$\|\Phi\| \leqslant A\sqrt{n}. \tag{6}$$

Si les nœuds sont racines d'un polynôme orthogonal de degré $n$, de poids borné $\omega(t) \geqslant c > 0$, alors

$$\|\Phi\| \leqslant An\ ^{*}). \tag{7}$$

Le problème aux limites considéré est équivalent à l'équation fonctionnelle

$$K_1 x \equiv Gx - \lambda Tx = y, \tag{8}$$

où

$$Gx = \frac{d^{2m}x}{dt^{2m}}, \quad z = Tx, \quad z(t) = \sum_{s=1}^{2m} p_s(t)\, x^{(2m-s)}(t). \tag{9}$$

L'inverse de l'opérateur $G$ est un opérateur intégral dont le noyau est la fonction de Green de l'opérateur différentiel $\frac{d^{2m}x}{dt^{2m}}$ sous les conditions (2).

L'existence de la fonction de Green résulte du fait que l'équation différentielle

$$\frac{d^{2m}x}{dt^{2m}} = 0 \tag{10}$$

---

*) Pour (5) et (7) voir Natanson [I], pages 539 et 544. La majoration (6) figure dans le livre de Szegö.

admet une solution unique vérifiant les conditions (2), la solution triviale, car la solution générale de l'équation (10) est un polynôme arbitraire de degré $2m - 1$, et tout polynôme différent du polynôme nul et vérifiant les conditions (2) doit comporter le facteur $(t-a)^m (b-t)^m$ de puissance $2m$.

Donc

$$x = G^{-1}y, \quad x(s) = \int_a^b g(s, t)\, y(t)\, dt \quad (x \in \mathrm{X},\ y \in \mathrm{Y}).$$

La fonction $g(s, t)$ possède des dérivées continues jusqu'à l'ordre $2m - 2$ et la dérivée d'ordre $2m - 1$ présente un saut pour $s = t$. D'où

$$x^{(k)}(s) = \int_a^b \frac{\partial^k}{\partial s^k} g(s, t)\, y(t)\, dt \quad (k = 0, 1, \ldots, 2m-1),$$

donc

$$\max_s |x^{(k)}(s)| \leqslant A_k \| y \|_{\mathrm{C}} \quad (k = 0, 1, \ldots, 2m-1). \tag{11}$$

Si l'on convient que $A_{2m} = 1$, alors (11) sera valable pour $k = 2m$.

Si l'on veut que l'opérateur $G$ réalise une isométrie entre X et Y, il faut que les normes de ces espaces soient telles que

$$\| x \|_{\mathrm{X}} = \| Gx \|_{\mathrm{Y}} = \max_t |x^{(2m)}(t)|.$$

De plus, (11) permet d'évaluer toute dérivée de la fonction $x$ au moyen de la norme de l'élément $x$ dans l'espace X:

$$\max_t |x^{(k)}(t)| \leqslant A_k \| x \| \quad (k = 0, 1, \ldots, 2m). \tag{12}$$

On remarquera que l'opérateur $G$ fait correspondre aux éléments de l'espace $\tilde{\mathrm{X}}$ ceux de l'espace $\tilde{\mathrm{Y}}$, c'est-à-dire des polynômes de degré $\leqslant n - 1$.

Evaluons la norme de $T$ comme opérateur de $\mathbf{C}^{(2m)}[a, b]$ dans $\mathbf{C}[a, b]$. Compte tenu de (12) on obtient

$$\| Tx \| = \max_t \left| \sum_{s=1}^{2m} p_s(t)\, x^{(2m-s)}(t) \right| \leqslant \left[ \sum_{s=1}^{2m} B_s A_{2m-s} \right] \| x \|,$$

où $B_s = \max_t |p_s(t)| \quad (s = 1, 2, \ldots, 2m)$.

Donc

$$\| T \| \leqslant \sum_{s=1}^{2m} B_s A_{2m-s}. \tag{13}$$

On évalue de façon analogue la norme de l'opérateur $K_1^{-1}$ si l'on connaît la fonction de Green de l'équation (1) sous les conditions (2).

L'équation approchée, le système (4), peut visiblement se mettre sous la forme

$$\widetilde{K}_1\widetilde{x} \equiv G\widetilde{x} - \lambda\Phi T\widetilde{x} = \Phi y, \qquad (14)$$

puisque le système (4) se déduit à partir de l'équation (1) par substitution de l'expression (3) de $\widetilde{x}$ et par identification des deux membres aux points $t_1, t_2, \ldots, t_n$, ce qui revient à identifier les polynômes d'interpolation construits d'après les valeurs prises en ces points.

En comparant l'équation (14) à l'équation (22) de 2.3, on constate qu'elle est construite de façon spéciale, si bien que la condition I ter est vérifiée pour $\mu = 0$.

Evaluons $\frac{d}{dt} Tx$ pour vérifier la condition II ter, c'est-à-dire pour établir le degré éventuel d'approximation des éléments de la forme $Tx$ par des éléments de $\widetilde{\mathrm{Y}}$. Compte tenu de (12) il vient

$$\left\| \frac{d}{dt} Tx \right\|_C = \max_t \left| \frac{d}{dt} \sum_{s=1}^{2m} p_s(t)\, x^{(2m-s)}(t) \right| =$$

$$= \max_t \left| \sum_{s=0}^{2m} (p_s'(t) + p_{s+1}(t))\, x^{(2m-s)}(t) \right| \leqslant M \, \|x\|$$

$$(p_0(t) = p_{2m+1}(t) = 0).$$

Le théorème de Jackson affirme l'existence d'un polynôme $\widetilde{y} \in \widetilde{\mathrm{Y}}$ tel que

$$\| Tx - \widetilde{y} \| \leqslant \frac{AM \, \| x \|}{n}.$$

Donc, la condition II ter est satisfaite pour $\mu_1 = O\left(\frac{1}{n}\right)$.

S'agissant de la condition III ter, si le second membre de l'équation (1), la fonction $y$, admet une dérivée continue, en appliquant le théorème de Jackson on obtient pour $\mu_2$ la valeur $O\left(\frac{1}{n}\right)$.

A noter que si dans l'équation (1) $p_1 = 0$ (ce qu'on peut toujours réaliser par un changement convenable des coordonnées) et que les coefficients $p_s$ soient 2-continûment dérivables, en effectuant les mêmes raisonnements sur $\frac{d^2}{dt^2} Tx$ on obtiendrait pour $\mu_1$ la valeur $O\left(\frac{1}{n^2}\right)$. Si $y$ est aussi 2-continûment dérivable, alors $\mu_2 = O\left(\frac{1}{n^2}\right)$.

Les résultats du § 2, les majorations de $\mu_1$ et $\mu_2$ et de $\| \Phi \|$ nous permettent de conclure d'après le théorème 2.1 ter que le système (4) admet une solution pour les $n$ assez grands.

Le théorème 2.3 ter établit la convergence des solutions approchées vers la solution exacte dans le cas de nœuds de Tchébychev ou de Gauss. De plus

$$\| x^* - \tilde{x}_n^* \| = O\left(\frac{\ln n}{n}\right), \quad \| x^* - \tilde{x}_n^* \| = O\left(\frac{1}{\sqrt{n}}\right)$$

respectivement pour les nœuds de Tchébychev et de Gauss.

Dans le cas déjà mentionné où $p_1 = 0$, le système (4) admet aussi une solution si pour nœuds on prend les zéros d'un polynôme de degré $n$, orthogonal par rapport au poids $\omega(t) \geqslant c > 0$. La majoration de l'erreur de la solution approchée est

$$\| x^* - \tilde{x}_n^* \| = O\left(\frac{1}{n}\right).$$

Si les coefficients sont suffisamment lisses, la vitesse de convergence peut devenir plus grande. Supposons que $p_1, p_2, \ldots, p_{2m}$ et $y$ sont $r$-dérivables et leurs dérivées $r$-ièmes vérifient la condition de Lipschitz dans le rapport $\alpha$. Utilisons la remarque qui suit le théorème 1.3 et qui dit que la vitesse de convergence est définie par

$$\| x^* - \tilde{x}_n^* \| = O(\varepsilon \| P \|) \quad (P = G^{-1}\Phi G, \ \| P \| = \| \Phi \|), \tag{15}$$

où $\varepsilon$ caractérise la possibilité d'approcher la solution $x^*$ par un élément $\tilde{x} \in \tilde{X}$. Comme

$$\| x^* - \tilde{x} \|_X = \| G(x^* - \tilde{x}) \|_Y = \left\| \frac{d^{2m}x^*}{dt^{2m}} - \tilde{y} \right\|_Y \quad (\tilde{y} = \tilde{G}x \in \tilde{Y}),$$

$\varepsilon$ est défini par le degré d'approximation de la dérivée $2m$-ième de la solution par un polynôme de degré $n - 1$. Or, on établit aisément par récurrence sur $r$ que la fonction $\frac{d^{2m}x^*}{dt^{2m}}$ admet une dérivée $r$-ième vérifiant la condition de Lipschitz dans le rapport $\alpha$. Donc, d'après le théorème de Jackson, il existe un polynôme $\tilde{y} \in \tilde{Y}$ tel que

$$\left\| \frac{d^{2m}x^*}{dt^{2m}} - \tilde{y} \right\| \leqslant \frac{D}{n^{r+\alpha}}.$$

Par suite, $\varepsilon$ est d'ordre $n^{-r-\alpha}$ et d'après (15)

$$\| x^* - \tilde{x}_n^* \| = O\left(\frac{\ln n}{n^{r+\alpha}}\right), \quad \| x^* - \tilde{x}_n^* \| = O\left(\frac{1}{n^{r+\alpha-\frac{1}{2}}}\right),$$

$$\| x^* - \tilde{x}_n^* \| = O\left(\frac{1}{n^{r+\alpha-1}}\right)$$

respectivement pour les nœuds de Tchébychev, Gauss et les zéros d'un polynôme orthogonal suivant un poids minoré.

REMARQUE 1. Sans modifier fondamentalement les raisonnements, on peut établir la convergence de cette méthode non seulement pour le problème aux limites, mais aussi pour le problème de Cauchy. La

nature des raisonnements n'est pas changée par la considération d'un système au lieu d'une seule équation différentielle.

REMARQUE 2. Le théorème de convergence de la méthode d'approximation considérée peut être traité comme un théorème de convergence d'un processus d'interpolation: le polynôme $\tilde{x}_n^*(t)$ peut être considéré comme un polynôme d'interpolation pour la solution $x^*(t)$, construit d'après les conditions aux limites et les valeurs de l'opérateur différentiel $Kx^*$ aux points $t_1, t_2, \ldots, t_n$.

REMARQUE 3. Il est curieux de noter que si les points sont équidistants, la méthode considérée ne converge pas même pour l'équation élémentaire $\frac{d^2x}{dt^2} = y$, ce qu'on établit sans peine en tenant compte de l'éventuelle divergence d'un processus d'interpolation ordinaire pour des nœuds équidistants (cf. Natanson [I], page 519).

5.2. Appliquons maintenant la théorie générale à l'étude de la convergence de la méthode de Galerkin et de la méthode des moments *).

Soient données une équation différentielle (1) et les conditions homogènes aux limites

$$M(x) = 0 \tag{16}$$

aux extrémités d'un intervalle. Le principe de la *méthode des moments* est le suivant: on prend un système de fonctions $\{\omega_k\}$ vérifiant les conditions aux limites (16) et on cherche la solution approchée sous la forme

$$\tilde{x}(t) = \sum_{k=1}^{n} c_k \omega_k(t), \tag{17}$$

les coefficients $c_k$ étant définis à partir de la condition d'orthogonalité du résultat de la substitution de $\tilde{x}$ dans l'équation (1) (ou plus exactement du résidu obtenu) aux $n$ premières fonctions d'un système $\{\zeta_j\}$:

$$\int_a^b L\left(\sum_{k=1}^{n} c_k \omega_k\right) \zeta_j(t)\,dt = \int_a^b y(t)\,\zeta_j(t)\,dt \quad (j = 1, 2, \ldots, n),$$

où $L(x)$ désigne

$$L(x) = \frac{d^{2m}x}{dt^{2m}} - \lambda \sum_{s=1}^{2m} p_s \frac{d^{2m-s}x}{dt^{2m-s}}.$$

Si $\zeta_j = \omega_j$, on obtient la *méthode de Galerkin*.

*) La convergence de la méthode de Galerkin dans sa forme générale a été établie pour la première fois dans un travail de M. Keldych [1]. Le cas particulier de la méthode de Galerkin, la méthode de Ritz, et la méthode des moments ont été traités antérieurement dans les travaux de N. Krylov et N. Bogolioubov (cf. Krylov ainsi que Mikhline [II]).

5.3. Soit en particulier l'équation

$$x''(t) - \lambda\,[p_1(t)\,x'(t) + p_2(t)\,x(t)] = y(t) \quad (|t| \leqslant 1) \tag{18}$$

avec les conditions

$$x(-1) = x(1) = 0. \tag{19}$$

Pour fonctions $\omega_k$ on prend

$$\omega_k(t) = t^{k-1}(1 - t^2) \quad (k = 1, 2, \ldots).$$

L'équation (18) peut encore s'écrire

$$Gx - \lambda Tx = y, \tag{20}$$

où $Gx = x''$. Pour X on prend l'espace unitaire (non complet) des fonctions 2-continûment dérivables, vérifiant la condition aux limites (19), dans lequel

$$\|x\|_X = \left[\int_{-1}^{1} |x''(t)|^2\,dt\right]^{1/2},$$

$$(x_1, x_2)_X = \int_{-1}^{1} x_1''(t)\,x_2''(t)\,dt.$$

Munissons Y, qui est constitué des fonctions continues, d'une métrique adaptée à celle de l'espace X par l'opérateur $G$:

$$\|y\|_Y = \|G^{-1}y\|_X = \left[\int_{-1}^{1} |y(t)|^2\,dt\right]^{1/2}, \quad (y_1, y_2)_Y = \int_{-1}^{1} y_1(t)\,y_2(t)\,dt.$$

Pour $\widetilde{X}$ on prend le sous-espace des fonctions de la forme (17) et on définit $\widetilde{Y}$ comme $G(\widetilde{X})$, ce qui, visiblement, donne l'ensemble des polynômes de degré $n - 1$. On définit enfin l'opérateur $\Phi$ comme un projecteur de Y sur $\widetilde{Y}$. Ceci étant, si $\zeta_1, \zeta_2, \ldots, \zeta_n$ est un système d'éléments linéairement indépendants de $\widetilde{Y}$, l'égalité $\Phi y = 0$ équivaut au système d'égalités $(y, \zeta_j) = 0$ $(j = 1, 2, \ldots, n)$.

Sous ces conditions, l'équation de la méthode des moments équivaut à l'équation fonctionnelle

$$G\tilde{x} - \lambda\Phi T\tilde{x} = \Phi y, \tag{21}$$

c'est-à-dire l'équation approchée est comme dans 2.3 la projection de l'équation exacte. Donc, la condition I ter est satisfaite pour $\mu = 0$.

Pour vérifier la condition II ter, il faut prouver qu'il est possible d'approcher l'élément $Tx$ par des éléments de $\widetilde{Y}$.

Notons préalablement que le maximum et la dérivée première d'une fonction de X peuvent être évalués à l'aide de la norme de cette

fonction:

$$|x'(t)|=\frac{1}{2}\left|\int_{-1}^{1}[x'(t)-x'(\tau)]\,d\tau\right|=\frac{1}{2}\left|\int_{-1}^{1}\left[\int_{\tau}^{t}x''(u)\,du\right]d\tau\right|\leqslant$$

$$\leqslant\frac{1}{2}\int_{-1}^{1}\left[\int_{\tau}^{t}|x''(u)|^2\,du\right]^{1/2}\left[\int_{\tau}^{t}du\right]^{1/2}d\tau\leqslant A\,\|x\|,$$

$$|x(t)|\leqslant\int_{-1}^{t}|x'(\tau)|\,d\tau\leqslant B\,\|x\|.$$

Si l'on admet que les coefficients de l'équation (18) sont continûment dérivables, on peut écrire

$$\left\|\frac{d}{dt}Tx\right\|_Y=\|p_1x''+(p_1+p_2')\,x'+p_2x\|_Y\leqslant C\,\|x\|.$$

D'où l'on déduit sans peine que la fonction $Tx$ vérifie la condition de Lipschitz dans le rapport 1/2

$$|(Tx)(t_1)-(Tx)(t_2)|=\left|\int_{t_1}^{t_2}\frac{d}{dt}Tx\,dt\right|\leqslant\left\|\frac{d}{dt}Tx\right\|_Y|t_2-t_1|^{1/2},$$

donc, elle peut être approchée par un polynôme $y\in Y$, si bien que

$$\|Tx-\tilde{y}\|\leqslant\frac{D\,\|x\|}{\sqrt{n}}.$$

Par suite, on peut admettre que $\mu_1=O\left(\frac{1}{\sqrt{n}}\right)$.

Des raisonnements analogues nous donnent pour $\mu_2$ la valeur $\mu_2=O\left(\frac{1}{\sqrt{n}}\right)$ (sous réserve que la dérivée du second membre soit de carré sommable).

Si $p_1=0$, on évalue de même $\frac{d^2}{dt^2}Tx$ et l'on obtient $\mu_1=$ $=O\left(\frac{1}{n\sqrt{n}}\right)$.

Le théorème 2.3 ter nous donne l'ordre de la vitesse de convergence des solutions approchées vers la solution exacte

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{\sqrt{n}}\right).$$

Si $p_1=0$ et $y''$ est de carré sommable, alors

$$\|x^*-\tilde{x}_n^*\|=O\left(\frac{1}{n\sqrt{n}}\right).$$

Si $p_1$, $p_2$ et $y$ possèdent des dérivées d'ordre plus élevé, on peut établir un ordre plus élevé de décroissance de l'erreur des solutions approchées, car, dans ce cas, $x^*$ est plusieurs fois différentiable et admet de ce fait une meilleure approximation par des polynômes.

La convergence des solutions approchées vers la solution exacte a lieu dans tous les cas dans l'espace X, c'est-à-dire les solutions approchées et leurs dérivées premières convergent *uniformément* vers la solution exacte et sa dérivée première respectivement. Les dérivées secondes des solutions approchées convergent en moyenne vers la dérivée seconde de la solution exacte.

5.4. Considérons l'équation (18) pour d'autres définitions de la norme dans les espaces X et Y. Plus exactement, convenons que

$$\|x\|_X = \max_{|t|\leqslant 1} |x''(t)|, \quad \|y\|_Y = \max_{|t|\leqslant 1} |y(t)| \quad (x \in X, \ y \in Y).$$

Les normes sont visiblement compatibles, si bien que $\|G\| = \|G^{-1}\| = 1$. Nous conservons la définition précédente de l'opérateur $\Phi$, ce qui nous permettra de traiter l'équation approchée comme la projection de l'équation exacte (cf. 2.3), donc $\mu = 0$. Evaluons la norme de $\Phi$. Soient $\zeta_1, \zeta_2, \ldots, \zeta_n$ des polynômes orthogonaux par rapport au poids de l'unité (polynômes de Legendre). La projection $\Phi$ de l'espace Y sur $\widetilde{Y}$ s'écrit

$$\Phi y = \sum_{k=1}^{n} (y, \zeta_k) \zeta_k.$$

Mais, en vertu d'une majoration classique du développement partiel d'une fonction continue en série suivant les polynômes de Legendre (cf. par exemple Agakhanov et Natanson [1]),

$$\max_{|t|\leqslant 1} \left| \sum_{k=1}^{n} (y, \zeta_k) \zeta_k(t) \right| \leqslant A \sqrt{n},$$

si bien que

$$\|\Phi\| \leqslant A \sqrt{n}. \tag{22}$$

Pour approcher l'élément $Tx$ par un polynôme, c'est-à-dire par un élément de $\widetilde{Y}$, on utilisera le fait que

$$\left\| \frac{d}{dt} Tx \right\|_Y \leqslant C \|x\|.$$

Donc, on peut trouver $\widetilde{y} \in \widetilde{Y}$ tel que

$$\|Tx - \widetilde{y}\| \leqslant \frac{D\|x\|}{n}.$$

Par suite, $\mu_1 = O\left(\frac{1}{n}\right)$ $\left(\text{si } p_1 = 0, \text{ on peut prendre } \mu_1 = O\left(\frac{1}{n^2}\right)\right)$. Enfin, $\mu_2 = \frac{E_n(y)}{\|y\|}$. Compte tenu de (22), le théorème 2.2 ter

nous donne

$$\| x^* - \tilde{x}_n^* \| = (\mu_1 + \mu_2) O(\sqrt{n}).$$

Si $E_n(y)\sqrt{n} \to 0$, les solutions approchées et leurs dérivées premières et secondes convergent uniformément respectivement vers la solution exacte et ses dérivées première et seconde. Si, de plus, $p_1 = 0$ et $y$ est 2-dérivable, alors

$$\| x^* - \tilde{x}_n^* \| = O\left(\frac{1}{n\sqrt{n}}\right).$$

5.5. Considérons maintenant l'équation

$$L(x) \equiv \frac{d}{dt}\left[p \frac{dx}{dt}\right] - \lambda \left\{\frac{d}{dt}[qx] + rx\right\} = y \tag{23}$$

avec les conditions

$$x(0) = x(1) = 0. \tag{24}$$

On admettra que les fonctions $p$ et $q$ sont continûment dérivables, $p(t) > 0$, et de plus que $\lambda$ est tel que le problème aux limites posé ait une solution unique.

Suivant la méthode de Galerkin, on cherche la solution sous la forme

$$\tilde{x}(t) = \sum_{k=1}^{n} c_k \omega_k(t) \quad (0 \leqslant t \leqslant 1), \tag{25}$$

où $\omega_k$ $(k = 1, 2, \ldots)$ sont des fonctions continûment dérivables, vérifiant la condition (24):

$$\omega_k(0) = \omega_k(1) = 0 \quad (k = 1, 2, \ldots)$$

et telles que leurs dérivées forment avec la fonction unité un système complet. On admettra que le système $\{\omega_k'\}$ est orthonormal par rapport au poids $p$:

$$\int_0^1 p(t)\, \omega_j'(t)\, \omega_k'(t)\, dt = \begin{cases} 0 & \text{pour } j \neq k, \\ 1 & \text{pour } j = k \end{cases} \quad (j, k = 1, 2, \ldots). \tag{26}$$

Ajoutons la fonction $\omega_0'(t) = \frac{1}{p(t)}\left[\int_0^1 \frac{ds}{p(s)}\right]^{-1/2}$ à ce système. La relation (26) aura lieu pour $j, k = 0$ et le système $\{\omega_k'\}$ $(k = 0, 1, 2, \ldots)$ sera un système orthonormal complet.

Traitons l'équation (23) avec les conditions (24) comme une seule équation fonctionnelle

$$Gx - \lambda Tx = y \tag{27}$$

dans l'espace X des fonctions 2-continûment dérivables vérifiant la condition aux limites (24). Définissons une norme dans X en posant

$$\| x \| = \left[ \int_0^1 p(t) \, | x'(t) |^2 \, dt \right]^{1/2}, \quad (x_1, x_2) = \int_0^1 p(t) \, x_1'(t) \, x_2'(t) \, dt.$$

Pour Y prenons l'espace des fonctions continues; la métrique de Y est choisie à partir de la condition de compatibilité avec celle de X:

$$\| y \|_Y = \| G^{-1} y \|_X \text{*)}, \quad (y_1, y_2)_Y = (G^{-1} y_1, G^{-1} y_2)_X.$$

Pour $\tilde{X}$ prenons l'ensemble des éléments de la forme (25) et pour $\tilde{Y}$ celui des éléments de la forme

$$y = \sum_{k=1}^{n} c_k G \omega_k. \tag{28}$$

Enfin, l'opérateur $\Phi$ est défini comme un projecteur de Y sur $\tilde{Y}$, si bien que $\| \Phi \| = 1$. L'égalité $\Phi y = 0$ équivaut à $(y, G\omega_k) = 0$ ($k = 1, 2, \ldots, n$), ce qui d'après la définition du produit scalaire dans Y équivaut à

$$(G^{-1} y, \omega_k)_X = \int_0^1 p(t) \frac{d}{dt} (G^{-1} y) \, \omega_k'(t) \, dt = - \int_0^1 y(t) \, \omega_k(t) \, dt = 0.$$

L'équation de la méthode de Galerkin

$$\int_0^1 [L(x)(t) - y(t)] \, \omega_j(t) \, dt = 0 \quad (j = 1, 2, \ldots, n)$$

ne signifie donc rien d'autre que

$$\Phi [Gx - \lambda Tx] = \Phi y,$$

c'est-à-dire l'équation approchée est la projection de l'équation exacte (cf. 2.3).

Pour appliquer les résultats du § 2 on a $\mu = 0$. Pour évaluer $\mu_1$ on se servira de la remarque de 2.1 (l'inégalité (6)), c'est-à-dire on approchera l'élément $z = G^{-1} Tx$, avec $\| x \| < 1$, par un élément de $\tilde{X}$ dans la métrique de X. On a $Gz = Tx$, ou

$$\frac{d}{dt} \left[ p \frac{dz}{dt} \right] = \frac{d}{dt} [qx] + rx.$$

*) L'élément $x = G^{-1} y$ est la solution de l'équation $\frac{d}{dt} \left( p \frac{dx}{dt} \right) = y$, qui vérifie les conditions (24).

Une intégration entre 0 et $u \leqslant 1$ nous donne

$$p(u)\,z'(u) = q(u)\,x(u) + \int_0^u r(t)\,x(t)\,dt + C =$$

$$= q(u)\int_0^u x'(t)\,dt + \int_0^u \left(\int_t^u r(\tau)\,d\tau\right) x'(t)\,dt + C =$$

$$= \int_0^1 h(u, t)\,x'(t)\,dt + C,$$

où

$$h(u, t) = \begin{cases} q(u) + \int_t^u r(\tau)\,d\tau & (t \leqslant u), \\ 0 & (t > u). \end{cases}$$

Pour éliminer $C$, intégrons l'égalité obtenue après avoir préalablement divisé les deux membres par $p(u)$:

$$0 = \int_0^1 \left[\int_0^1 \frac{1}{p(u)} h(t, u)\,du\right] x'(t)\,dt + C\int_0^1 \frac{du}{p(u)}.$$

En tirant $C$ de cette relation et en le portant dans l'égalité précédente on obtient

$$z'(u) = \int_0^1 K(u, t)\,x'(t)\,dt,$$

où

$$K(u, t) = \frac{1}{p(u)} h(u, t) - \frac{1}{p(u)\int_0^1 \frac{ds}{p(s)}} \int_0^1 \frac{1}{p(s)} h(s, t)\,ds.$$

On remarquera que de toute évidence

$$\int_0^1 K(u, t)\,du = 0 \quad (0 \leqslant t \leqslant 1).$$

Le noyau $K(u, t)$ est approximable par un développement partiel en série de Fourier suivant $\{\omega_j'(u)\}$ $(j = 1, 2, \ldots)$:

$$K_n(u, t) = \sum_{j=1}^n \alpha_j(t)\,\omega_j'(u) \quad \left(\alpha_j(t) = \int_0^1 p(u)\,K(u, t)\,\omega_j'(u)\,du\right),$$

puisque le noyau est orthogonal à la fonction $1/p\,(u)$ et le système $\{\omega_j'\}$ $(j = 0, 1, \ldots)$ complet. Donc

$$\int_0^1 p\,(u)\,[K\,(u,\,t)-K_n\,(u,\,t)]^2\,du \underset{n\to\infty}{\longrightarrow} 0 \quad (0\leqslant t\leqslant 1) \tag{29}$$

de façon monotone.

En déterminant $\tilde{x}$ à partir de

$$\tilde{x}'\,(u)=\int_0^1 K_n\,(u,\,t)\,x'\,(t)\,dt=\sum_{j=1}^{n} c_j\omega_j'\,(u) \quad \left(c_j=\int_0^1 \alpha_j\,(t)\,x'\,(t)\,dt\right),$$

on obtient l'approximation souhaitée. En effet, on a

$$\|z-\tilde{x}\|^2=\int_0^1 p\,(u)\,|\,z'\,(u)-\tilde{x}'\,(u)\,|^2\,du=$$

$$=\int_0^1 p\,(u)\left|\int_0^1 [K\,(u,\,t)-K_n\,(u,\,t)]\,x'\,(t)\,dt\right|^2 du\leqslant$$

$$\leqslant\int_0^1 p\,(t)\,|\,x'\,(t)\,|^2\,dt\int_0^1\left[\int_0^1 |\,K\,(u,\,t)-K_n\,(u,\,t)\,|^2\,p\,(u)\,du\right]\frac{dt}{p\,(t)}=$$

$$=\eta_n^2\,\|\,x\,\|^2,$$

où

$$\eta_n^2=\int_0^1 \frac{dt}{p\,(t)}\int_0^1 [K\,(u,\,t)-K_n\,(u,\,t)]^2\,p\,(u)\,du \underset{n\to\infty}{\longrightarrow} 0.$$

On établit de façon analogue que $\mu_2$ est petit. Il faut prouver que l'élément $z = G^{-1}y$ se laisse approximer par un élément de $\tilde{X}$. On a

$$\frac{d}{dt}\left(p\,\frac{dz}{dt}\right)=y, \quad p\,\frac{dz}{dt}=\int y\,dt=F,$$

$F$ étant choisi tel que $\int_0^1 \frac{F\,(t)}{p\,(t)}\,dt=0$, donc $\frac{dz}{dt}=\frac{F}{p}$ peut être approchée par une somme de la forme $\sum_{k=1}^{n} c_k\omega_k'$.

Si $\omega_k$ sont des fonctions trigonométriques, on peut évaluer l'ordre d'approximation. Plus exactement, soit

$$\omega_k\,(t) = \sin k\pi t \quad (k = 1, 2, \ldots).$$

Il faut approcher le noyau $K(u, t)$ par des sommes de la forme

$$\sum_{k=1}^{n} \alpha_k(t) \cos k\pi u,$$

où

$$\alpha_k(t) = 2 \int_0^1 K(u, t) \cos k\pi u \, du$$

sont d'ordre $1/k$ en tant que coefficients de Fourier d'une fonction à variation bornée, et l'écart

$$\int_0^1 [K(u, t) - K_n(u, t)]^2 du = 2 \sum_{k=n+1}^{\infty} |\alpha_k(t)|^2 = O\left(\frac{1}{n}\right).$$

Donc, $\mu_1 = O\left(\frac{1}{\sqrt{n}}\right)$. On vérifie immédiatement que $\mu_2 = O\left(\frac{1}{\sqrt{n}}\right)$.

Donc, dans le cas étudié, le théorème 2.3 ter nous conduit à la majoration

$$\| x^* - \tilde{x}_n^* \| = O\left(\frac{1}{\sqrt{n}}\right),$$

c'est-à-dire les solutions approchées convergent vers la solution exacte pour la métrique de l'espace X, ce qui, en particulier, entraîne la convergence uniforme avec la même vitesse.

Si l'expression différentielle $L(x)$ est auto-adjointe, c'est-à-dire si $q = 0$, la méthode de Galerkin coïncide avec celle de Ritz. Si l'on admet alors que la dérivée première par rapport à $u$ de la fonction $p(u) K(u, t)$ est une fonction à variation bornée (ce sera le cas si les dérivées de $p$ et $r$ le sont aussi), on obtient l'ordre $1/k^2$ pour les coefficients de Fourier, ce qui nous amène à la majoration

$$\mu_1 = O\left(\frac{1}{n\sqrt{n}}\right)$$

et à la majoration respective de la vitesse de convergence des solutions approchées vers la solution exacte.

L'application de la remarque 1 qui suit le théorème 1.3 au cas d'une bonne approximabilité de la solution par les fonctions du système permet d'obtenir une plus grande vitesse de convergence des solutions approchées.

## § 6. Application aux problèmes aux limites pour équations de type elliptique

**6.1.** Nous allons appliquer ici la théorie générale des méthodes d'approximation à la résolution approchée des problèmes aux limites pour certaines équations de type elliptique. Nous considérons

notamment le problème aux limites pour l'équation

$$\Delta u + \lambda a u = v \tag{1}$$

dans un domaine $D$ limité par une courbe $\Gamma$, avec la condition aux limites

$$u\,|_\Gamma = 0 \tag{2}$$

en dimension $m$ (= 2, 3). Pour mener à bien notre étude, outre la théorie générale, nous aurons besoin de certains résultats de physique mathématique et de théorie de l'approximation que nous citerons sans les démontrer.

Théorème de Gunter-Korn. *Si $u$ est solution de l'équation*

$$\Delta u = f, \quad u\,|_\Gamma = 0, \tag{3}$$

*où $\Gamma$ est un contour lisse et la fonction $f$ vérifie la condition de Lipschitz dans le rapport $\beta$ et avec la constante $M$ ($f \in \mathrm{Lip}_M\,\beta$), alors $u\,(x,y)$ possède des dérivées partielles secondes vérifiant la condition de Lipschitz dans n'importe quel rapport $\beta' < \beta$, et de plus*

$$\frac{\partial^2 u}{\partial x^2},\ \frac{\partial^2 u}{\partial x\,\partial y},\ \frac{\partial^2 u}{\partial y^2} \in \mathrm{Lip}_{M'}\beta' \quad (M' \leqslant CM, \quad C = C(\beta, \beta'))\ *).$$

On sait que si les fonctions d'une variable jouissent de certaines propriétés de dérivabilité, il est possible de les approximer, en même temps que leurs dérivées, par des polynômes trigonométriques ou algébriques. Ceci étant, si la fonction s'annule aux extrémités de $[a, b]$, on peut faire en sorte que les fonctions approximantes s'y annulent aussi, et notamment exiger dans le cas d'une approximation par des polynômes algébriques qu'elles renferment le facteur $(t - a)\,(b - t)$.

Les fonctions de plusieurs variables sont justiciables de théorèmes analogues. Nous en citons deux : l'un établissant la complétude du système de fonctions correspondant, l'autre caractérisant la précision de l'approximation.

Théorème de complétude **). *Soit donné un domaine $D$ limité par un contour $\Gamma$, d'équation*

$$\omega\,(x, y) = 0,$$

*où $\omega$ est une fonction continûment différentiable par morceaux, et de plus*

$$\omega\,(x, y) > 0 \ \ ((x, y) \in D), \ \mathrm{grad}\ \omega\,(x, y) \neq 0 \ \ ((x, y) \in \Gamma). \tag{4}$$

*Alors le système de fonctions de la forme*

$$\tilde{u}\,(x, y) = \omega\,(x, y)\,P\,(x, y) \quad (P \text{ est un polynôme})$$

*) Cf. Gunter [1].
**) Cf. Kantorovitch et Krylov, page 296.

*est complet dans l'espace* $\overset{\circ}{W}{}_2^{(1)}$ *(c'est-à-dire dans l'espace des fonctions de* $W_2^{(1)}$ *s'annulant sur* $\Gamma$*).*

THÉORÈME DE HARRIK *). *Supposons qu'une fonction* $\omega$ *remplit les conditions du théorème de complétude, mais qu'en plus elle est k-continûment dérivable et ses dérivées k-ièmes vérifient la condition de Lipschitz. Si une fonction u est k-dérivable et ses dérivées k-ièmes vérifient la condition de Lipschitz dans le rapport* $\alpha$*, il existe une suite de polynômes* $P_n$ *de degré* $\leqslant n$*, tels que les fonctions* $\omega P_n$ *approximent u avec ses dérivées, plus précisément*

$$\| u - \omega P_n \|_{C^{(r)}} = O\left(\frac{1}{n^{k+\alpha-r}}\right), \tag{5}$$

*où*

$$\| u \|_{C^{(r)}} = \sum_{i=0}^{r} \sum_{\nu_1+\nu_2=i} \max_D \left| \frac{\partial^i u}{\partial x^{\nu_1} \partial y^{\nu_2}} \right|. \tag{6}$$

6.2. Voyons maintenant le problème aux limites posé au début du paragraphe. On cherchera la solution approchée sous la forme

$$\tilde{u}(x, y) = \omega(x, y) P(x, y), \tag{7}$$

où $P$ est un polynôme de degré $\leqslant n$, et $\omega$ vérifie les conditions du théorème de Harrik pour $k = 1$.

Les coefficients du polynôme $P$ (en nombre $N$) sont déterminés à partir du système d'équations

$$\iint_D (\Delta\tilde{u} + \lambda a\tilde{u}) \zeta_k \, dx\, dy = \iint_D v\zeta_k \, dx\, dy \quad (k = 1, 2, \ldots, N) \tag{8}$$

(on parlera plus bas du choix de $\zeta_k$).

Considérons l'espace $\mathbf{U}$ des fonctions $u$ 2-continûment dérivables s'annulant sur $\Gamma$, muni de la métrique et du produit scalaire par

$$\| u \| = \left[ \iint_D |\Delta u|^2 \, dx\, dy \right]^{1/2}, \quad (u_1, u_2) = \iint_D \Delta u_1 \Delta u_2 \, dx\, dy.$$

Prenons:

pour $\tilde{\mathbf{U}}$ l'ensemble des fonctions $\tilde{u}$;

pour $\mathbf{V}$ l'espace des fonctions de la forme $v = \Delta u$ ($u \in \mathbf{U}$), autrement dit des $v$ tels que la solution de l'équation $\Delta u = v$ appartienne à $\mathbf{U}$;

pour métrique de $\mathbf{V}$ la métrique de l'espace hilbertien $\mathbf{L}^2(D)$;

pour $\tilde{\mathbf{V}}$ enfin, l'ensemble des fonctions de la forme $\tilde{v} = \Delta\tilde{u}$ ($\tilde{u} \in \tilde{\mathbf{U}}$).

*) Cf. Harrik [2].

Ceci posé, on peut traiter le problème aux limites comme une équation fonctionnelle de la forme

$$Gu + \lambda Tu = v \quad (Gu = \Delta u, \quad Tu = au), \tag{9}$$

les métriques des espaces **U** et **V** étant adaptées de telle sorte que $G$ réalise une isométrie entre **U** et **V**.

Pour $\zeta_1, \zeta_2, \ldots, \zeta_N$ on peut prendre tout système linéairement indépendant de fonctions dans $\widetilde{\mathbf{V}}$, par exemple

$$\zeta_{ij} = \Delta\,[\omega(x, y)\, x^i y^j] \quad (i + j \leqslant n).$$

Si par $\Phi$ on entend la projection orthogonale de **V** sur $\widetilde{\mathbf{V}}$, alors l'équation $\Phi v = 0$ équivaut au système d'équations

$$(v, \zeta_k) = \iint_D v\zeta_k\, dx\, dy = 0 \quad (k = 1, 2, \ldots, n).$$

Le système (8) s'écrit donc dans les nouvelles notations

$$\Phi G\tilde{u} - \Phi T\tilde{u} = \Phi v, \tag{10}$$

c'est-à-dire l'équation approchée est la projection de l'équation exacte, donc la condition I ter est satisfaite pour $\mu = 0$.

Pour vérifier la condition II ter, nous allons montrer que pour $u \in \mathbf{U}$ on a $Tu = au \in \text{Lip}\,\beta$, où $\beta < 1$ si $m = 2$ et $\beta < 1/2$ si $m = 3$.

En effet, $u \in \mathbf{W}_2^{(2)}$ puisque $u \in \mathbf{U}$. D'après la remarque suivant le théorème d'immersion (XI.4.4, tome 1), $u \in \text{Lip}\,\beta$, où

$$\beta < 1 \quad (m = 2), \quad \beta < 1/2 \quad (m = 3).$$

Il s'ensuit, en vertu du théorème de Gunter-Korn, que les dérivées secondes de $z = G^{-1}Tu = \Delta^{-1}au$ appartiennent à Lip $\beta'$, où $\beta' < \beta$, c'est-à-dire $\beta' < 1$ pour $m = 2$ et $\beta' < 1/2$ pour $m = 3$, et de plus

$$\left\| \frac{\partial^2 z}{\partial x^2} \right\|_{\text{Lip}\,\beta'} \leqslant C \| u \|.$$

D'après le théorème de Harrik, la fonction $z$ et ses dérivées admettent une approximation d'ordre $\frac{1}{n^{\beta'}}$ par des fonctions de la forme $\tilde{u}$, d'où la majoration

$$\| G^{-1}Tu - \tilde{u} \| \leqslant \frac{C_1}{n^{\beta'}}.$$

Donc, la condition II ter est vérifiée pour $\mu = O\,(1/n^{\beta'})$.

De façon analogue, si $v$ appartient à Lip $\beta$, notamment si $v \in \mathbf{W}_2^{(2)}$, le même raisonnement montre que $\mu_2 = O\,(1/n^{\beta'})$. Tout ce qui vient d'être dit permet de conclure dans ce cas à la convergence de la méthode des moments pour les principales fonctions des formes

indiquées, cette convergence ayant lieu pour la métrique de U:

$$\| u^* - \tilde{u}_n^* \|_U = O\left(\frac{1}{n^{\beta'}}\right).$$

D'où l'on conclut immédiatement à la convergence uniforme des fonctions. En effet, en se servant de la représentation de $u$ au moyen de la fonction de Green, on obtient

$$| u^*(x, y) - \tilde{u}_n^*(x, y) | = \left| \iint_D g(x, y)\, \Delta (u^* - \tilde{u}_n^*)\, dx\, dy \right| \leqslant$$
$$\leqslant \left[ \iint_D | g(x, y) |^2\, dx\, dy \right]^{1/2} \| u^* - \tilde{u}_n^* \| = O\left(\frac{1}{n^{\beta'}}\right).$$

A noter que si l'on dispose de plus de données sur les propriétés différentielles de la solution, on peut établir une plus grande vitesse de convergence.

6.3. La convergence de la méthode de Galerkin peut être établie sans indication de l'ordre exact de la vitesse dans des conditions bien plus générales que pour la méthode des moments.

Considérons le problème aux limites pour l'équation

$$\Delta u + \lambda \left( au + b \frac{\partial u}{\partial x} + c \frac{\partial u}{\partial y} \right) = v \tag{11}$$

avec la condition

$$u|_\Gamma = 0. \tag{12}$$

On admettra que les coefficients $a$, $b$ et $c$ sont continûment dérivables.

Cherchons la solution sous la même forme que dans 6.2. Déterminons les coefficients à partir du système de la méthode de Galerkin

$$\iint_D \left[ \Delta \tilde{u} + \lambda \left( a\tilde{u} + b \frac{\partial \tilde{u}}{\partial x} + c \frac{\partial \tilde{u}}{\partial y} \right) \right] \zeta_k\, dx\, dy = \iint_D v \zeta_k\, dx\, dy$$
$$(k = 1, 2, \ldots, N).$$

Les fonctions $\zeta_k$ sont ici de la forme $\zeta_k = \omega(x, y) P_k(x, y)$, où $P_k(x, y)$ sont des polynômes de degré $\leqslant n$. Munissons U d'une métrique

$$(u_1, u_2) = - \iint_D u_1 \Delta u_2\, dx\, dy = \iint_D \left[ \frac{\partial u_1}{\partial x} \frac{\partial u_2}{\partial x} + \frac{\partial u_1}{\partial y} \frac{\partial u_2}{\partial y} \right] dx\, dy,$$

$$\| u \|^2 = - \iint_D u \Delta u\, dx\, dy = \iint_D \left[ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 \right] dx\, dy,$$

et V respectivement

$$(v_1, v_2)_V = (G^{-1} v_1, G^{-1} v_2)_U, \quad \| v \|_V = \| G^{-1} v \|_U,$$

c'est-à-dire si $u_1$ et $u_2$ sont des solutions des équations $\Delta u = v_1$ et $\Delta u = v_2$, appartenant à **U**, alors on convient que

$$(v_1, v_2) = -\iint_D u_1 v_2 \, dx\, dy = -\iint_D u_2 v_1 \, dx\, dy,$$

$$\| v_1 \|^2 = -\iint_D u_1 v_1 \, dx\, dy.$$

Le problème aux limites envisagé peut être mis sous forme d'une seule équation fonctionnelle dans l'espace **U**:

$$Gu + \lambda Tu = v \quad \left(Gu = \Delta u, \quad Tu = au + b\frac{\partial u}{\partial x} + c\frac{\partial u}{\partial y}\right),$$

et l'équation approchée devient

$$G\tilde{u} + \lambda \Phi T \tilde{u} = \Phi v.$$

Pour mettre en application le théorème 1.6 vérifions la compacité de l'opérateur $G^{-1}T$.

Il est immédiat de voir que l'espace **U** est linéairement isométrique à une partie dense de l'espace $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$. Pour le prouver, il suffit de prendre pour métrique de $\mathbf{W}_2^{(1)}$ la métrique définie par la fonctionnelle $\int_\Gamma u\,(x, y)\, d\Gamma$,

$$\|u\|_{\mathbf{W}_2^{(1)}} = \left| \int_\Gamma u\,(x, y)\, d\Gamma \right| + \left\{ \iint_D |\operatorname{grad} u|^2 \, dx\, dy \right\}^{1/2} =$$

$$= \left\{ \iint_D \left[ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right] dx\, dy \right\}^{1/2} = \| u \|_{\mathbf{U}},$$

donc l'opérateur $T$ peut être considéré comme un opérateur linéaire continu de $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ dans $\mathbf{L}^2$. L'inégalité de Bernstein-Ladyjenskaïa nous dit que l'opérateur inverse $\Delta^{-1}$ de $\mathbf{L}^2$ dans $\overset{\circ}{\mathbf{W}}{}_2^{(2)}$ est continu (cf. Ladyjenskaïa, chap. II, § 6). Donc, $\Delta^{-1}T \in B(\overset{\circ}{\mathbf{W}}{}_2^{(1)}, \overset{\circ}{\mathbf{W}}{}_2^{(2)})$. L'opérateur d'immersion de $\overset{\circ}{\mathbf{W}}{}_2^{(2)}$ dans $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ est compact (XI.3.5, tome 1), donc $\Delta^{-1}T$ regardé comme un opérateur de $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ dans $\overset{\circ}{\mathbf{W}}{}_2^{(2)}$ est également compact.

D'après le théorème de complétude, la suite des projections $P_n = G^{-1}\Phi G$ converge vers l'opérateur identité sur **U**. Donc, on peut se servir du théorème 1.6 qui dit que la méthode de Galerkin est con-

vergente dans l'espace U. D'où l'on déduit que les solutions approchées convergent uniformément vers la solution exacte.

A noter que les raisonnements précédents sont valables sans aucun changement au cas où le nombre de variables indépendantes est supérieur à deux.

Indiquons en conclusion que d'autres applications intéressantes des méthodes exposées dans ce chapitre sont accessibles dans de nombreux travaux et notamment dans Kalandia [1], Vladimirov [1], Philippov et Riabenki.

CHAPITRE XV

# MÉTHODE DE LA PLUS FORTE PENTE

La méthode de la plus forte pente (on dit encore de la pente maximum) est actuellement l'une des méthodes de résolution des problèmes d'extrémum libre les plus répandues. A l'origine elle fut considérée comme une méthode variationnelle de résolution des équations fonctionnelles linéaires et de détermination des valeurs propres d'opérateurs linéaires. Comme dans toute méthode variationnelle, la résolution de l'équation (la détermination des valeurs propres) se ramène à un problème d'extrémum d'une fonctionnelle de forme spéciale, définie sur l'espace tout entier. Cette méthode convient à la minimisation de fonctionnelles d'une forme bien plus générale que celles mentionnées ci-dessus. En un nombre de pas peu élevé, elle conduit à un voisinage assez petit d'un point stationnaire (au voisinage d'un point minimum si la fonctionnelle est convexe).

L'idée de la méthode est la suivante : on construit une suite d'approximations du minimum telle que le passage d'une approximation à la suivante se fasse suivant la direction de la plus forte décroissance de la fonctionnelle considérée. D'où la dénomination de la méthode.

**L'idée de cette méthode remonte à Cauchy qui l'a considérée dans le cas d'un espace de dimension finie. Dans le cas général de fonctionnelles quadratiques, cette méthode a été développée par L. Kantorovitch (cf. Kantorovitch [7], [9]). Pour la démonstration des théorèmes généraux énoncés ci-après (§§ 1 et 2) on s'est inspiré des travaux de M. Birman [1], [2]. Certaines modifications de cette méthode et ses applications à divers problèmes sont accessibles dans Vaïnberg; Krasnosselski et autres; Lioubitch et Maïstrovski [1]. L. Zah a dans sa thèse (1955) appliqué la méthode de la plus forte pente à la résolution d'équations différentielles elliptiques (§ 3). Au § 4 on utilise une approche qui a déjà servi à l'étude de la méthode du gradient lié (cf. Démianov et Rubinov).**

## § 1. Résolution des équations linéaires

**1.1.** Soit $\Phi$ une fonctionnelle réelle (non linéaire) définie sur un espace X normé (réel ou complexe). Supposons que $\Phi$ est minorée sur X et cherchons l'élément $x^* \in X$ (s'il existe) qui la minimise :

$$\Phi(x) \geqslant \Phi(x^*) \quad (x \in X).$$

Pour résoudre ce problème on procède généralement comme suit: on construit une suite $\{x_n\}$ « minimisant » la fonctionnelle $\Phi$ au sens que

$$\lim_{n\to\infty} \Phi(x_n) = \inf_{x\in X} \Phi(x).$$

Dans certains cas il est possible de construire une suite $\{x_n\}$ convergente vers un élément $x^*$. Si l'on admet que $\Phi$ est continue, cet élément sera solution du problème posé.

La méthode de la plus forte pente est une méthode de construction d'une suite $\{x_n\}$ qui est minimisante pour une assez vaste classe de fonctionnelles. Cette méthode peut être énoncée pour des fonctionnelles $\Phi$ dérivables suivant des directions.

Soit $x$ un élément fixe de X. Considérons la demi-droite issue de $x$ dans la direction de $z$, c'est-à-dire l'ensemble des éléments de la forme $x + \alpha z$, où $\alpha$ est un nombre réel positif, $z \neq 0$ un élément de X caractérisant la direction de la demi-droite *). La restriction de la fonctionnelle $\Phi$ à la demi-droite indiquée est une fonction de variable réelle que l'on désignera dans la suite par $\varphi(\alpha; x, z)$. Plus exactement

$$\varphi(\alpha; x, z) = \Phi(x + \alpha z) \quad (\alpha \geqslant 0).$$

On conçoit que la *dérivée* de la fonctionnelle $\Phi$ suivant la direction de $z$ (au point $x$) est par définition l'expression

$$\frac{\partial\Phi}{\partial z}(x) = \frac{1}{\|z\|}\varphi'(\alpha; x, z)\Big|_{\alpha=0} = \frac{1}{\|z\|}\lim_{h\to 0+}\frac{\Phi(x+hz)-\Phi(x)}{h}.$$

On admettra que la dérivée $\frac{\partial\Phi}{\partial z}(x)$ existe pour chaque direction et que, de plus, il existe une direction suivant laquelle cette dérivée est minimale. Cette direction s'appelle direction de la plus forte pente de la fonctionnelle $\Phi$ en $x$.

Passons maintenant à la description de la méthode. Supposons que la fonctionnelle $\Phi$ est dérivable en chaque point $x$ suivant toutes les directions, et de plus pour chaque $x \in X$ il existe une direction de plus forte pente. Choisissons arbitrairement un élément $x_0 \in X$ (l'approximation nulle vers le minimum). Supposons $x_k$ trouvé. Il semble naturel de chercher l'approximation suivante sous la forme

$$x_{k+1} = x_k + \varepsilon_{k+1} z_{k+1}.$$

$z_{k+1}$ est la direction de plus forte pente au point $x_k$. Le paramètre numérique $\varepsilon_{k+1}$ est appelé valeur de la pente. On peut le déterminer à partir de considérations différentes. Indiquons trois d'entre elles.

*) Deux éléments non nuls $z_1$ et $z_2$ définissent une même direction si, et seulement si, ils sont positivement proportionnels (c'est-à-dire $z_1 = \lambda z_2$ pour un $\lambda > 0$), ou ce qui revient au même s'ils sont situés sur une même demi-droite issue de l'origine. D'une façon générale, la direction est définie par la donnée d'un vecteur portée par la demi-droite.

1) Supposons que $\frac{\partial \Phi}{\partial z_{k+1}}(x) < 0$ et que la dérivée $\frac{\partial \Phi}{\partial z}(x_k + \alpha z_{k+1})$ est continue comme fonction de $\alpha$. Alors la fonction $\varphi(\alpha;\ x_k,\ z_{k+1})$ est décroissante au moins sur l'intervalle $[0,\ \alpha'_k]$, où $\alpha'_k$ est la plus petite racine positive de l'équation $\varphi'(\alpha;\ x_k,\ z_{k+1}) = 0$. Cette racine est la valeur de la pente $\varepsilon_{k+1}$.

2) Supposons que la fonction $\varphi(\alpha;\ x_k,\ z_{k+1})$ atteint son minimum en un point du demi-axe positif. Ce point est la valeur de la pente.

Si $\Phi$ est une fonctionnelle strictement convexe, les deux méthodes indiquées coïncident (pour plus de détails voir 4.4).

3) Donnons-nous une suite de nombres strictement positifs $\{\varepsilon_k\}$ telle que $\varepsilon_k \to 0$ et $\sum_{k=0}^{\infty} \varepsilon_k = \infty$ (on admet que la direction de $z_k$ est normée, c'est-à-dire $\| z_k \| = 1$). Nous admettons que la valeur de la pente est égale à $\varepsilon_k$ au $k$-ième pas.

Après avoir opté pour un procédé de détermination de la valeur de la pente, on obtient une suite $\{x_k\}$ qui est minimisante dans de nombreux cas. La construction décrite constitue le principe de la méthode de la plus forte pente.

La marche à suivre est la même si l'on cherche le maximum d'une fonctionnelle majorée. Dans ce cas il faut considérer la direction de plus forte montée.

1.2. La minimisation d'une fonctionnelle et la résolution d'une équation fonctionnelle linéaire sont reliées de la manière suivante.

Soit $U$ un opérateur auto-adjoint dans un espace hilbertien **H**. Désignons comme toujours par $m$ et $M$ les bornes inférieure et supérieure de $U$ (cf. IX.4.3, tome 1). Supposons que $m > 0$ et considérons l'équation

$$Ux = y. \tag{1}$$

L'opérateur inverse $U^{-1}$ existe, puisque $\lambda = 0$ n'appartient pas au spectre de $U$ (théorème IX.5.3, tome 1), donc l'équation (1) possède une solution unique $x^* = U^{-1}y$ quel que soit l'élément $y \in \mathbf{H}$.

Formons la fonctionnelle

$$F(x) = (Ux,\ x) - [(x,\ y) + (y,\ x)] \quad (x \in \mathbf{H}). \tag{2}$$

Théorème 1. *La solution $x^*$ de l'équation* (1) *minimise la fonctionnelle* (2). *Inversement, si un élément $\tilde{x}$ minimise la fonctionnelle* (2), *il est solution de l'équation* (1), *c'est-à-dire $\tilde{x} = x^*$.*

Démonstration. Comme $y = Ux^*$, on peut mettre la fonctionnelle $F$ sous la forme

$$F(x) = (Ux,\ x) - (x,\ Ux^*) - (Ux^*,\ x) = \\ = (U(x - x^*),\ x - x^*) - (Ux^*,\ x^*). \tag{3}$$

D'où

$$F(x) \geqslant m(x - x^*, x - x^*) - (Ux^*, x^*) \geqslant -(Ux^*, x^*) = F(x^*),$$

c'est-à-dire $x^*$ minimise la fonctionnelle $F$.

Cette inégalité entraîne également la deuxième partie du théorème, puisque

$$0 = F(\tilde{x}) - F(x^*) = (U(\tilde{x} - x^*), \tilde{x} - x^*) \geqslant m(\tilde{x} - x^*, \tilde{x} - x^*),$$

donc $\tilde{x} = x^*$.

Remarque. Si l'existence de la solution $x^*$ est établie, on voit de la démonstration du théorème que la condition $m > 0$ peut être remplacée par une condition plus faible: $(Ux, x) > 0$ pour tout $x \neq 0$.

**1.3.** Appliquée à la fonctionnelle (2), la méthode de la plus forte pente conduit à une suite convergeant vers $x^*$, solution de l'équation (1), ce qui permet de la considérer comme une méthode de résolution approchée de l'équation (1).

La marche à suivre est la suivante.

Si $x, z \in \mathbf{H}$, alors

$$F(x + z) = (U(x + z), x + z) - [(x + z, y) + (y, x + z)] =$$
$$= (Ux, x) - [(x, y) + (y, x)] + [(Ux - y, z) + (z, Ux - y)] +$$
$$+ (Uz, z) = F(x) + [(Ux - y, z) + (z, Ux - y)] + (Uz, z), \quad (4)$$

d'où

$$\frac{\partial F}{\partial z}(x) = \frac{1}{\| z \|}[(Ux - y, z) + (z, Ux - y)] = \frac{2\,\mathrm{Re}\,(Ux - y, z)}{\| z \|}.$$

Donc, la direction de plus forte pente au point $x_0$ est définie par l'élément $(-z_1)$, où

$$z_1 = Ux_0 - y.$$

Pour déterminer $\varepsilon_1$ formons l'équation

$$\varphi'(\alpha; x_0, -z_1) = 0,$$

où, en vertu de (4),

$$\varphi(\alpha; x_0, -z_1) = F(x_0 - \alpha z_1) = F(x_0) - 2\alpha(z_1, z_1) +$$
$$+ \alpha^2 (Uz_1, z_1).$$

Donc

$$\varepsilon_1 = \frac{(z_1, z_1)}{(Uz_1, z_1)}.$$

En définitive

$$x_1 = x_0 - \varepsilon_1 z_1.$$

De façon analogue

$$x_2 = x_1 - \varepsilon_2 z_2 \quad \left(z_2 = Ux_1 - y, \quad \varepsilon_2 = \frac{(z_2, z_2)}{(Uz_2, z_2)}\right)$$

et en général

$$x_n = x_{n-1} - \varepsilon_n z_n \quad (n = 1, 2, \ldots),$$

où

$$z_n = Ux_{n-1} - y, \quad \varepsilon_n = \frac{(z_n, z_n)}{(Uz_n, z_n)} \quad (n = 1, 2, \ldots).$$

La suite $\{x_n\}$ est minimisante. Bien plus, on a le

Théorème 2. *La suite $\{x_n\}$ converge vers $x^*$, la vitesse de convergence étant donnée par*

$$\| x_n - x^* \| \leqslant C \left( \frac{M-m}{M+m} \right)^n \quad \left( n = 0, 1, \ldots; \quad C = \frac{\| z_1 \|}{m} \right). \tag{5}$$

Démonstration. Mettons l'équation (1) sous la forme

$$x = x - kUx + ky. \tag{6}$$

Choisissons le coefficient $k > 0$ tel que l'opérateur

$$T = I - kU$$

soit de norme minimale. La norme sera minimale si

$$1 - km = -(1 - kM),$$

où $1 - km$ est la borne supérieure et $1 - kM$ la borne inférieure de $T$. D'où

$$k = \frac{2}{M+m}.$$

Ceci étant

$$\| T \| = 1 - km = kM - 1 = \frac{M-m}{M+m}. \tag{7}$$

Désignons

$$\tilde{x}_1 = Tx_0 + ky = x_0 - k(Ux_0 - y) = x_0 - kz_1. \tag{8}$$

Considérons maintenant l'opérateur $V = U^{1/2}$ (cf. théorème V.6.2, tome 1). Grâce à (3) on peut mettre la fonctionnelle $F$ sous la forme

$$F(x) = (U(x - x^*), x - x^*) - (Ux^*, x^*) = (V(x - x^*), V(x - x^*)) -$$

$$- (Vx^*, Vx^*) = \| V(x - x^*) \|^2 - \| Vx^* \|^2. \tag{9}$$

L'inégalité évidente $F(\tilde{x}_1) \geqslant F(x_1)$ entraîne

$$F(x_1) - F(x^*) \leqslant F(\tilde{x}_1) - F(x^*),$$

ce qu'on peut encore écrire grâce à (9)

$$\| V(x_1 - x^*) \| \leqslant \| V(\tilde{x}_1 - x^*) \|. \tag{10}$$

L'équation (6) est équivalente à l'équation (1). Donc

$$x^* = Tx^* + ky.$$

En retranchant ceci de (8), on trouve

$$\tilde{x}_1 - x^* = T(x_0 - x^*)$$

et par suite

$$V(\tilde{x}_1 - x^*) = TV(x_0 - x^*),$$

d'où

$$\| V(\tilde{x}_1 - x^*) \| \leqslant \| T \| \, \| V(x_0 - x^*) \| = \frac{M-m}{M+m} \| V(x_0 - x^*) \|.$$

En comparant avec (10) on obtient

$$\| V(x_1 - x^*) \| \leqslant \frac{M-m}{M+m} \| V(x_0 - x^*) \|. \qquad (11)$$

Les mêmes raisonnements pour chaque $n = 1, 2, \ldots$ conduisent aux inégalités

$$\| V(x_n - x^*) \| \leqslant \frac{M-m}{M+m} \| V(x_{n-1} - x^*) \| \quad (n = 1, 2, \ldots).$$

Et en définitive

$$\| V(x_n - x^*) \| \leqslant \left(\frac{M-m}{M+m}\right)^n \| V(x_0 - x^*) \| \quad (n = 1, 2, \ldots). \qquad (12)$$

La fonction $t^{-1/2}$ étant continue sur $[m, M]$ et *a fortiori* sur le spectre de l'opérateur $U$, l'opérateur inverse $V^{-1}$ existe, et de plus (cf. théorème IX.5.2, tome 1)

$$\| V^{-1} \| = \max_{t \in S_U} \frac{1}{\sqrt{t}} = \frac{1}{\sqrt{m}}.$$

A noter encore que

$$\| V \| = \max_{t \in S_U} \sqrt{t} = \sqrt{M}.$$

D'après ce qui précède, il vient compte tenu de (12)

$$\| x_n - x^* \| = \| V^{-1} V(x_n - x^*) \| \leqslant \| V^{-1} \| \left(\frac{M-m}{M+m}\right)^n \| V(x_0 - x^*) \| \leqslant$$
$$\leqslant \| V^{-1} \| \, \| V \| \left(\frac{M-m}{M+m}\right)^n \| x_0 - x^* \| = \sqrt{\frac{M}{m}} \left(\frac{M-m}{M+m}\right)^n \| x_0 - x^* \|.$$

Cette majoration est défavorable, car le second membre contient l'élément inconnu $x^*$. Pour éviter cela on remarquera que

$$\| V(x_0 - x^*) \| = \| V^{-1} U(x_0 - x^*) \| \leqslant \| V^{-1} \| \, \| Ux_0 - y \| = \frac{\| z_1 \|}{\sqrt{m}},$$

d'où

$$\| x_n - x^* \| \leqslant \frac{\| z_1 \|}{m} \left( \frac{M-m}{M+m} \right)^n \quad (n = 1, 2, \ldots),$$

ce qui est confondu avec (5).

C.q.f.d.

**1.4.** Exprimons l'approximation $x_p$ en fonction de $x_0$ et $z_1$. On a

$$x_2 = x_1 - \varepsilon_2 z_2 = x_1 - \varepsilon_2 (Ux_1 - y) =$$
$$= x_0 - \varepsilon_1 z_1 - \varepsilon_2 (Ux_0 - y - \varepsilon_1 Uz_1) = x_0 - (\varepsilon_1 + \varepsilon_2) z_1 +$$
$$+ \varepsilon_1 \varepsilon_2 Uz_1 = x_0 + \lambda_1^{(2)} z_1 + \lambda_2^{(2)} Uz_1.$$

On prouve sans peine par récurrence que

$$x_p = x_0 + \lambda_1^{(p)} z_1 + \ldots + \lambda_p^{(p)} U^{p-1} z_1. \tag{13}$$

Considérons l'élément

$$x_{(1)} = x_0 + \lambda_1 z_1 + \ldots + \lambda_p U^{p-1} z_1$$

et déterminons les coefficients complexes $\lambda_1, \lambda_2, \ldots, \lambda_p$ à partir de la condition de minimalité de la fonctionnelle $F(x_{(1)})$. Comme

$$F(x_{(1)}) = (Ux_0 + \sum_{j=1}^{p} \lambda_j U^j z_1, \ x_0 + \sum_{j=1}^{p} \lambda_j U^{j-1} z_1) -$$
$$- [(x_{(1)}, y) + (y, x_{(1)})] = F(x_0) + \sum_{j=1}^{p} (\lambda_j + \overline{\lambda}_j)(U^{j-1} z_1, z_2) +$$
$$+ \sum_{j=1}^{p} \sum_{k=1}^{p} \lambda_j \overline{\lambda}_k (U^j z_1, U^{k-1} z_1),$$

en désignant $\lambda_k = \sigma_k + \tau_k i$ $(k = 1, 2, \ldots, p)$ on obtient pour la détermination de $\lambda_k$ deux systèmes d'équations

$$\frac{\partial F(x_{(1)})}{\partial \sigma_j} = 0, \quad \frac{\partial F(x_{(1)})}{\partial \tau_j} = 0 \quad (j = 1, 2, \ldots, p),$$

ou, sous la forme détaillée

$$(U^{j-1} z_1, z_1) + \sum_{k=1}^{p} \sigma_k (U^j z_1, U^{k-1} z_1) = 0, \quad \sum_{k=1}^{p} \tau_k (U^j z_1, U^{k-1} z_1) = 0.$$

En multipliant la deuxième égalité par $i$, en ajoutant à la première et en tenant compte du fait que $(U^j z_1, U^{k-1} z_1) = (U^{j+k-1} z_1, z_1)$, on peut grouper les deux systèmes en un seul

$$(U^{j-1} z_1, z_2) + \sum_{k=1}^{p} \lambda_k (U^{j+k-1} z_1, z_1) = 0 \quad (j = 1, 2, \ldots, p). \tag{14}$$

Comme le minimum cherché existe manifestement, ce système admet une solution *).

A partir de l'élément $x_{(1)}$ construisons de façon analogue un élément $x_{(2)}$ et ainsi de suite. On obtient en définitive une suite $x_{(0)} = x_0, x_{(1)}, \ldots$ correspondant à la *variante en p pas* de la méthode de la plus forte pente.

De la construction de la suite du processus en $p$ pas, il est clair que la convergence vers $x^*$ sera au moins $p$ fois plus rapide que pour la suite $\{x_n\}$. En effet, en se servant de la démonstration du théorème 2, on établit sans peine que

$$\| x_{(n)} - x^* \| \leqslant \frac{\| z_1 \|}{m} \left[ \left( \frac{M-m}{M+m} \right)^p \right]^n. \tag{15}$$

Mais, en fait la convergence est bien plus rapide.

THÉORÈME 3 (Birman). *On a la majoration*

$$\| x_{(n)} - x^* \| \leqslant \frac{\| z_1 \|}{m} \left[ \frac{1}{\vartheta_p \left( \frac{M+m}{M-m} \right)} \right]^n \quad (n = 1, 2, \ldots), \tag{16}$$

*où* $\vartheta_p(t)$ *est un polynôme de Tchébychev de degré* $p$:

$$\vartheta_p(t) = \cos p \arccos t \quad (t \in [-1, 1]).$$

DÉMONSTRATION. Soit l'équation

$$x = Tx + \tilde{y} \quad (T = I - U\varphi(U), \ \tilde{y} = \varphi(U) y), \tag{17}$$

où $\varphi(t)$ est un polynôme de degré $p - 1$.

Désignons

$$\tilde{x}_1 = Tx_0 + \tilde{y} = x_0 - \varphi(U)(Ux_0 - y) = x_0 - \varphi(U) z_1 = x_0 + \sum_{k=1}^{p} c_k U^{k-1} z_1.$$

Il est clair que $F(\tilde{x}_1) \geqslant F(x_{(1)})$. En reprenant les raisonnements du théorème 2, on obtient

$$\| x_{(n)} - x^* \| \leqslant \frac{\| z_1 \|}{m} \| T \|^n \quad (n = 1, 2, \ldots). \tag{18}$$

Comme

$$\| T \| \leqslant \max_{t \in [m, M]} | 1 - t\varphi(t) |,$$

on conçoit qu'il faut déterminer le polynôme $\varphi(t)$ à partir de la condition de minimalité de $\max\limits_{t \in [m, M]} | 1 - t\varphi(t) |$. Si l'on désigne $\psi(t) = 1 - t\varphi(t)$, alors $\psi(t)$ sera le polynôme de degré $p$ qui s'écar-

*) Cette solution n'est pas forcément unique. Cependant la valeur de $F(x_{(1)})$ ne dépend pas du choix de la solution.

tera le moins de zéro sur l'intervalle $[m, M]$ et de plus $\psi(0) = 1$. On sait que *)

$$\psi(t) = \frac{\vartheta_p\left(\frac{2t-M-m}{M-m}\right)}{\vartheta_p\left(\frac{M+m}{M-m}\right)}.$$

Comme

$$\max_{t\in[m,\,M]}\left|\vartheta_p\left(\frac{2t-M-m}{M-m}\right)\right| = \max_{t\in[-1,\,1]}|\vartheta_p(t)| = 1,$$

on obtient pour $\|T\|$ la majoration

$$\|T\| \leqslant \max_{t\in[m,\,M]}|\psi(t)| = \frac{1}{\vartheta_p\left(\frac{M+m}{M-m}\right)},$$

ce qui, en vertu de (18), prouve le théorème.

REMARQUE 1. Soit $\rho(t)$ un polynôme de degré $p$. On a alors (Natanson [I], page 78)

$$|\rho(t_0)| \leqslant |\vartheta_p(t_0)| \max_{t\in[-1,\,1]}|\rho(t)| \quad (|t_0| > 1).$$

En prenant $\rho(t) = t^p$, $t_0 = \frac{M+m}{M-m}$, on aura

$$\frac{1}{\vartheta_p\left(\frac{M+m}{M-m}\right)} \leqslant \left(\frac{M-m}{M+m}\right)^p.$$

Donc, (15) est une conséquence de (16).

REMARQUE 2. Comme

$$\frac{\vartheta_p(t)}{t^p} \xrightarrow[t\to\infty]{} \frac{1}{2^{p-1}},$$

la majoration (16) diffère de (15) d'un facteur proche de $\left(\frac{1}{2^{p-1}}\right)^n$ si $\frac{M+m}{M-m}$ est assez grand.

Samokich [1] a obtenu une meilleure majoration que (16) sous des hypothèses supplémentaires sur le spectre de l'opérateur $U$.

REMARQUE 3. Pour $m = 0$, on démontre que si l'équation (1) admet une solution (pas nécessairement unique), la méthode de la plus forte pente converge vers la solution $x^*$ la plus proche de $x_0$ (Friedman [1]). En outre, on a la majoration $F(x_n) - F(x^*) = O(1/n)$, mais la vitesse de convergence de $x_n$ vers $x^*$ peut être aussi petite que l'on veut.

*) Cf. Gontcharov [1], page 230.

## § 2. Détermination des valeurs propres des opérateurs compacts

**2.1.** Considérons maintenant un opérateur auto-adjoint compact $U$. On supposera sans nuire à la généralité que $M > |m| \geqslant 0$ (cf. IX.4.3, tome 1). Alors

$$M = \sup_{\|x\|=1} (Ux, x) = \sup_{x \neq 0} \frac{(Ux, x)}{(x, x)} = \frac{(Ux^*, x^*)}{(x^*, x^*)},$$

où $x^*$ est l'élément propre associé à la plus grande valeur propre $\lambda_1 = M$. Donc, $\lambda_1$ est la plus grande valeur de la fonctionnelle

$$L(x) = \frac{(Ux, x)}{(x, x)}$$

et, en vertu de IX.4.3 (tome 1), tout élément réalisant le maximum est élément propre.

Utilisons la méthode de la plus forte pente *) pour déterminer le maximum de la fonctionnelle $L$. Soit un élément normé $x_0 \in \mathbf{H}$. Comme

$$L(x_0 + \alpha z) = \frac{(Ux_0, x_0) + \alpha(Ux_0, z) + \alpha(z, Ux_0) + \alpha^2(Uz, z)}{1 + \alpha(x_0, z) + \alpha(z, x_0) + \alpha^2(z, z)}, \tag{1}$$

alors

$$\frac{\partial L}{\partial z}(x_0) = \frac{[(Ux_0, z) + (z, Ux_0)] - [(x_0, z) + (z, x_0)](Ux_0, x_0)}{\|z\|} =$$
$$= \frac{(Ux_0 - \mu_0 x_0, z) + (z, Ux_0 - \mu_0 x_0)}{\|z\|},$$

où

$$\mu_0 = L(x_0) = (Ux_0, x_0).$$

De l'expression $\frac{\partial L}{\partial z}(x_0)$ il est clair que la direction du « gradient » est donnée par l'élément $z_1 = Ux_0 - \mu_0 x_0$ (on admet que $z_1 \neq 0$).

Donc, $\varepsilon_1$ se déduit de l'équation

$$\varphi'(\alpha; x_0, z_1) = 0, \tag{2}$$

où

$$\varphi(\alpha; x_0, z_1) = L(x_0 + \alpha z_1).$$

Comme

$$(x_0, z_1) = (x_0, Ux_0 - \mu x_0) = (x_0, Ux_0) - \mu_0(x_0, x_0) = 0,$$

$$(Ux_0, z_1) = (\mu_0 x_0 + z_1, z_1) = \mu_0(x_0, z_1) + (z_1, z_1) = (z_1, z_1),$$

des transformations élémentaires nous permettent de mettre l'équation (2) sous la forme

$$\frac{2\{(z_1, z_1) + [(Uz_1, z_1) - \mu_0(z_1, z_1)]\alpha - (z_1, z_1)^2\alpha^2\}}{[1 + \alpha^2(z_1, z_1)]^2} = 0.$$

*) Plus exactement, la méthode de « la plus forte montée ».

Donc, $\varepsilon_1$ est racine positive de l'équation

$$(z_1, z_1)^2\alpha^2 - [(Uz_1, z_1) - \mu_0 (z_1, z_1)]\, \alpha - (z_1, z_1) = 0,$$

c'est-à-dire

$$\varepsilon_1 = \frac{[(Uz_1, z_1) - \mu_0 (z_1, z_1)] + \sqrt{[(Uz_1, z_1) - \mu_0 (z_1, z_1)]^2 + 4 (z_1, z_1)^3}}{2 (z_1, z_1)^2}.$$

Posons

$$x^{(1)} = x_0 + \varepsilon_1 z_1$$

et prenons pour approximation suivante l'élément

$$x_1 = \frac{x^{(1)}}{\| x^{(1)} \|} = \frac{x_0 + \varepsilon_1 z_1}{\| x_0 + \varepsilon_1 z_1 \|}.$$

En effectuant la même procédure avec l'élément $x_1$ et les éléments suivants on obtient la suite $x_0, x_1, \ldots, x_n, \ldots$

$$x_n = \frac{x^{(n)}}{\| x^n \|}, \quad x^{(n)} = x_{n-1} + \varepsilon_n z_n \quad (n = 1, 2, \ldots),$$

où $(n = 1, 2, \ldots)$

$$z_n = Ux_{n-1} - \mu_{n-1} x_{n-1}, \quad \mu_{n-1} = L (x_{n-1}) = (Ux_{n-1}, x_{n-1}),$$

$$\varepsilon_n = \frac{[(Uz_n, z_n) - \mu_{n-1} (z_n, z_n)] + \sqrt{[(Uz_n, z_n) - \mu_{n-1} (z_n, z_n)]^2 + 4 (z_n, z_n)^3}}{2 (z_n, z_n)^2}.$$

2.2. S'agissant de la convergence de la suite $\{\mu_n\}$ vers $\lambda_1$ et de la suite $\{x_n\}$ vers l'élément propre correspondant, on a le

THÉORÈME 1. *Si un élément $x_0$ n'est pas orthogonal au sous-espace propre associé à la plus grande valeur propre $\lambda_1 = M$, alors $\mu_n \to \lambda_1$ et $x_n \to x^*$, où $x^*$ est l'élément propre associé à la valeur propre $\lambda_1$.*

DÉMONSTRATION. Désignons par $\lambda_k$ $(k = 1, 2, \ldots)$ les valeurs propres de l'opérateur $U$, par $\mathbf{H}_{\lambda_k} = \mathbf{H}_k$ les sous-espaces propres correspondants et par $P_k$ les projecteurs sur ces sous-espaces. Posons d'autre part

$$x_k^* = \frac{P_k x_0}{\| P_k x_0 \|} \; {}^{*)}.$$

D'après IX.4.5 (tome 1), on a

$$x_0 = \sum_k P_k x_0 = \sum_k \| P_k x_0 \| x_k^* = \sum_k c_{0k} x_k^*$$

$$(c_{0k} = \| P_k x_0 \|, \; k = 1, 2, \ldots). \quad (3)$$

Il est clair que $c_{0k} \geqslant 0$ $(k = 1, 2, \ldots)$ et, par hypothèse, $c_{01} > 0$. Les termes du second membre de (3) étant deux à deux orthogonaux

*) Si $P_k x_0 = 0$, alors $x_k^* = 0$.

et $\| x_k^* \| = 1$, il vient

$$\sum_k c_{0k}^2 = \| x_0 \|^2 = 1$$

et enfin

$$\mu_0 = (Ux_0, x_0) = \sum_k \lambda_k c_{0k}^2 \quad (\mu_0 \leqslant \lambda_1).$$

L'élément $x_1$ peut aussi être exprimé en fonction de $x_1^*$, $x_2^*$, ... En effet

$$x_1 = \frac{x^{(1)}}{\| x^{(1)} \|} = \frac{x_0 + \varepsilon_1 z_1}{\| x^{(1)} \|} = \frac{x_0 + \varepsilon_1 (Ux_0 - \mu_0 x_0)}{\| x^{(1)} \|}.$$

Or

$$Ux_0 = \sum \lambda_k c_{0k} x_k^*,$$

donc

$$x_1 = \sum_k c_{1k} x_k^* \quad \left( c_{1k} = \frac{[1 + \varepsilon_1 (\lambda_k - \mu_0)] c_{0k}}{\| x^{(1)} \|};\ k = 1, 2, \ldots \right).$$

Comme précédemment

$$\sum_k c_{1k}^2 = 1, \quad \mu_1 = \sum_k \lambda_k c_{1k}^2, \quad c_{11} > 0.$$

De façon générale

$$x_n = \sum_k c_{nk} x_k^*$$

$$\left( c_{nk} = \frac{[1 + \varepsilon_n (\lambda_k - \mu_{n-1})] c_{n-1k}}{\| x^{(n)} \|};\ k, n = 1, 2, \ldots \right), \tag{4}$$

en outre

$$\sum_k c_{nk}^2 = 1, \quad \mu_n = \sum_k \lambda_k c_{nk}^2, \quad c_{n1} > 0 \quad (n = 0, 1, \ldots). \tag{5}$$

Exprimons $\mu_n$ en fonction de $\mu_{n-1}$. En vertu de la formule (1), on a

$$\mu_n = L(x_n) = L(x^{(n)}) = L(x_{n-1} + \varepsilon_n z_n) =$$
$$= \frac{(Ux_{n-1}, x_{n-1}) + 2\varepsilon_n (Ux_{n-1}, z_n) + \varepsilon_n^2 (Uz_n, z_n)}{1 + 2\varepsilon_n (x_{n-1}, z_n) + \varepsilon_n^2 (z_n, z_n)}.$$

Mais

$$(x_{n-1}, z_n) = (x_{n-1}, Ux_{n-1} - \mu_{n-1} x_{n-1}) = 0,$$
$$(Ux_{n-1}, z_n) = (\mu_{n-1} x_{n-1} + z_n, z_n) = (z_n, z_n).$$

Donc

$$\mu_n = \frac{\mu_{n-1} + 2\varepsilon_n (z_n, z_n) + \varepsilon_n^2 (Uz_n, z_n)}{1 + \varepsilon_n^2 (z_n, z_n)} =$$
$$= \mu_{n-1} + \frac{2\varepsilon_n (z_n, z_n) + \varepsilon_n^2 [(Uz_n, z_n) - \mu_{n-1} (z_n, z_n)]}{1 + \varepsilon_n^2 (z_n, z_n)}.$$

De plus $\varepsilon_n$ est solution de l'équation

$$(z_n,\ z_n)^2\,\varepsilon_n^2 - [(Uz_n,\ z_n) - \mu_{n-1}(z_n,\ z_n)]\,\varepsilon_n - (z_n,\ z_n) = 0, \qquad (6)$$

de sorte que

$$\varepsilon_n^2\,[(Uz_n,\ z_n) - \mu_{n-1}(z_n,\ z_n)] = \varepsilon_n^3\,(z_n,\ z_n)^2 - \varepsilon_n\,(z_n,\ z_n).$$

En tenant compte de ceci, on obtient

$$\mu_n = \mu_{n-1} + \frac{2\varepsilon_n\,(z_n,\ z_n) + \varepsilon_n^3\,(z_n,\ z_n)^2 - \varepsilon_n\,(z_n,\ z_n)}{1 + \varepsilon_n^2\,(z_n,\ z_n)},$$

d'où, en définitive,

$$\mu_n = \mu_{n-1} + \varepsilon_n\,(z_n,\ z_n). \qquad (7)$$

Comme $\varepsilon_n > 0$, il vient $\mu_n > \mu_{n-1}$, donc étant bornée la suite $\{\mu_n\}$ $(\mu_n \leqslant \lambda_1)$ admet une limite que nous désignons par $\mu$:

$$\mu = \lim_{n\to\infty} \mu_n.$$

De (7) il suit que

$$\lim_{n\to\infty} \varepsilon_n\,(z_n,\ z_n) = 0. \qquad (8)$$

L'inégalité

$$|\ \varepsilon_n\ (Uz_n,\ z_n)\ | \leqslant \|\ U\ \|\ \varepsilon_n\ (z_n,\ z_n)$$

entraîne

$$\lim_{n\to\infty} \varepsilon_n\,(Uz_n,\ z_n) = 0. \qquad (9)$$

En tenant compte de (8) et (9), on trouve en passant à la limite dans l'équation (6)

$$\lim_{n\to\infty} (z_n,\ z_n) = 0.$$

Donc

$$\lim_{n\to\infty} [Ux_n - \mu_n x_n] = \lim_{n\to\infty} z_{n+1} = 0. \qquad (10)$$

L'opérateur $U$ étant compact et la suite $\{x_n\}$ bornée, il existe une sous-suite $\{x_{n_j}\}$ telle que $\{Ux_{n_j}\}$ soit convergente. De la relation (10), il résulte que la suite $\{x_{n_j}\}$ est convergente. Supposons que $x_{n_j} \to \tilde{x}$. En passant à la limite dans l'égalité

$$Ux_{n_j} - \mu_{n_j} x_{n_j} = z_{n_j+1},$$

on trouve

$$U\tilde{x} - \mu\tilde{x} = 0.$$

Donc, $\mu$ est valeur propre et $\tilde{x}$ l'élément propre qui lui correspond. Montrons que $\mu = \lambda$ et $\tilde{x} = x_1^*$. Supposons que $\mu = \lambda_s$ $(s > 1)$.

Compte tenu de (4), il vient

$$\tilde{x} = \sum_k c_k x_k^* \quad (c_k = \lim_{j\to\infty} c_{n_j k}\,; \quad k = 1,\ 2,\ \ldots).$$

Or, $(\tilde{x},\ x_k^*) = 0$ $(k \neq s)$, donc $\tilde{x} = c_s x_s^*$, et de plus $|c_s| = 1$. Comme $c_{ns} \geqslant 0$, on trouve finalement $c_s = 1$ et $\tilde{x} = x_s^*$.

Considérons le rapport $\frac{c_{ns}}{c_{n1}}$. D'après (4), il vient

$$\frac{c_{ns}}{c_{n1}} = \frac{1+\varepsilon_n(\lambda_s - \mu_{n-1})}{1+\varepsilon_n(\lambda_1 - \mu_{n-1})}\,\frac{c_{n-1,\,s}}{c_{n-1,\,1}} < \frac{c_{n-1,\,s}}{c_{n-1,\,1}},$$

c'est-à-dire ce rapport décroît lorsque $n$ croît. Mais $\lim\limits_{j\to\infty} c_{n_j s} =$ $= c_s = 1$, donc $\lim\limits_{j\to\infty} c_{n_j 1} = c_1 = 0$ et par suite

$$\lim_{j\to\infty} \frac{c_{n_j s}}{c_{n_j 1}} = \infty,$$

ce qui est contradictoire. Donc $\mu = \lambda_1$, $\tilde{x} = x_1^*$.

Prouvons maintenant la convergence de la suite $\{x_n\}$ vers $x_1^*$. Soit $d$ la distance du point $\lambda_1$ au reste du spectre de $U$. On a

$$\| x_n - x_1^* \|^2 = \| x_n \|^2 + \| x_1^* \|^2 - (x_n,\ x_1^*) - (x_1^*,\ x_n) =$$
$$= 2\,(1 - c_{n1}) \leqslant 2\,(1 - c_{n1}^2) = 2 \sum_{k\geqslant 2} c_{nk}^2 \leqslant \frac{2}{d} \sum_k (\lambda_1 - \lambda_k)\, c_{nk}^2.$$

Comme en vertu de (23)

$$\sum_k (\lambda_1 - \lambda_k)\, c_{nk}^2 = \lambda_1 \sum_k c_{nk}^2 - \sum_k \lambda_k c_{nk}^2 = \lambda_1 - \mu_n,$$

il vient

$$\| x_n - x_1^* \| \leqslant \frac{2}{d}(\lambda_1 - \mu_n) \quad (n = 0,\ 1,\ \ldots). \tag{11}$$

D'où il suit que $x_n \to x_1^*$, puisque $\mu_n \to \lambda_1$.

C.q.f.d.

Remarque. S'agissant de la vitesse de convergence de $\mu_n \to \lambda_1$ on peut établir que

$$\lambda_1 - \mu_n \leqslant q^2\,(1 + \alpha_n)\,(\lambda_1 - \mu_{n-1}) \quad (n = 1,\ 2,\ \ldots),$$

où

$$q = \frac{M - m - d}{M - m + d},$$

et que la suite $\{\alpha_n\}$ tend vers zéro en décroissant de façon monotone.

Si donc $n$ est assez grand pour que $q\,(1 + \alpha_n) < 1$, la vitesse de convergence est caractérisée par une progression géométrique.

La vitesse de convergence de $x_n \to x_1^*$ peut être caractérisée à l'aide de la majoration (11).

Dans Samokich [2], la méthode de la plus forte pente est appliquée à la recherche des valeurs propres.

## § 3. Application aux équations différentielles elliptiques

**3.1.** Appliquons la méthode de la plus forte pente à la résolution d'une équation différentielle elliptique. Pour simplifier, nous considérons le problème aux limites pour l'équation auto-adjointe du second ordre avec deux variables indépendantes dans un disque fermé $D$ (centré en l'origine des coordonnées) limité par le cercle $S$:

$$Lx \equiv -\frac{\partial}{\partial s}\left(a\frac{\partial x}{\partial s}\right)-\frac{\partial}{\partial t}\left(b\frac{\partial x}{\partial t}\right)+cx=\varphi, \quad x|_S=0; \tag{1}$$

les coefficients $a$ et $b$ étant supposés être des fonctions continûment différentiables; les fonctions $c$ et $\varphi$, continues; $a,\ b>0$; $c\geqslant 0$ dans le disque $D$.

Appliquons l'opérateur $-\Delta$ aux deux membres de cette équation. On obtient une nouvelle équation

$$Ux \equiv -\Delta^{-1}Lx = \psi \ (\psi = -\Delta^{-1}\varphi), \tag{2}$$

que l'on résoudra dans le sous-espace $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ des fonctions de $\mathbf{W}_2^{(1)}$ s'annulant sur $S$. On rappelle (cf. XIV.6.3) que l'opérateur $\Delta^{-1}$ est un opérateur linéaire continu de $\mathbf{L}^2$ dans $\overset{\circ}{\mathbf{W}}{}_2^{(2)}$, si bien que $\psi\in\overset{\circ}{\mathbf{W}}{}_2^{(2)}$.

Vérifions que l'opérateur $U=-\Delta^{-1}L$ réalise les conditions du § 1. Soit $x$ une fonction de $\overset{\circ}{\mathbf{W}}{}_2^{(1)}$ 2-continûment dérivable. $Lx$ étant une fonction continue, on a $y=Ux=-\Delta^{-1}Lx\in\overset{\circ}{\mathbf{W}}{}_2^{(2)}\subset\overset{\circ}{\mathbf{W}}{}_2^{(1)}$. En appliquant deux fois la formule de Green on obtient

$$\begin{aligned}(Ux,\ x)&=(-\Delta^{-1}Lx,\ x)=(y,\ x)=\\ &=\iint_D\left(\frac{\partial y}{\partial s}\frac{\partial x}{\partial s}+\frac{\partial y}{\partial t}\frac{\partial x}{\partial t}\right)ds\,dt=-\iint_D x\left[\frac{\partial^2 y}{\partial s^2}+\frac{\partial^2 y}{\partial t^2}\right]ds\,dt=\\ &=-\iint_D x\,\Delta y\,ds\,dt=\iint_D x\left[-\frac{\partial}{\partial s}\left(a\frac{\partial x}{\partial s}\right)-\frac{\partial}{\partial t}\left(b\frac{\partial x}{\partial t}\right)+cx\right]ds\,dt=\\ &=\iint_D\left[a\left(\frac{\partial x}{\partial s}\right)^2+b\left(\frac{\partial x}{\partial t}\right)^2+cx^2\right]ds\,dt.\end{aligned} \tag{3}$$

Cette relation donne

$$(Ux,\ x)\geqslant\alpha\iint_D\left[\left(\frac{\partial x}{\partial s}\right)^2+\left(\frac{\partial x}{\partial t}\right)^2\right]ds\,dt=\alpha\,(x,\ x), \tag{4}$$

où

$$\alpha = \min [\min a(s, t), \min b(s, t)] \quad ((s, t) \in D).$$

Le théorème d'immersion (théorème XI.4.3, tome 1) nous dit que

$$\iint_D |x(s, t)|^2 \, ds\, dt \leqslant A \, \| x \|^2_{\overset{\circ}{W}{}_2^{(1)}}.$$

En utilisant ceci dans (3), on trouve

$$(Ux, x) \leqslant \beta \iint_D \left[ \left( \frac{\partial x}{\partial s} \right)^2 + \left( \frac{\partial x}{\partial t} \right)^2 \right] ds\, dt = \beta (x, x), \tag{5}$$

où

$$\beta = 2 \max [\max a(s, t), \max b(s, t), A \max c(s, t)] \quad ((s, t) \in D).$$

En raisonnant comme dans la déduction de (3), on est conduit à l'égalité

$$(Ux, y) = (x, Uy), \tag{6}$$

qui est valable pour toutes fonctions $x, y \in \overset{\circ}{W}{}_2^{(1)}$ 2-continûment différentiables. En combinant (4), (5) et (6) on trouve que l'opérateur $U$ considéré sur l'ensemble des fonctions 2-continûment différentiables de $\overset{\circ}{W}{}_2^{(1)}$ est borné, donc il le sera dans $\overset{\circ}{W}{}_2^{(1)}$, puisque cet ensemble de fonctions est dense dans $\overset{\circ}{W}{}_2^{(1)}$. Les relations (4,) (5) et (6) se généralisent en outre à l'espace $\overset{\circ}{W}{}_2^{(1)}$ tout entier. Donc, $U$ est un opérateur auto-adjoint dont les bornes inférieure et supérieure sont limitées par

$$m \geqslant \alpha > 0, \quad M \leqslant \beta. \tag{7}$$

Appliquons la méthode de la plus forte pente à l'équation (2). Si $x_0 \in \overset{\circ}{W}{}_2^{(1)}$ est une approximation initiale, les approximations successives $x_n$ se déterminent par les formules de XV.1.3

$$x_n = x_{n-1} - \varepsilon_n z_n \quad (n = 1, 2, \ldots),$$

où

$$z_n = Ux_{n-1} - \psi = \Delta^{-1} (Lx_{n-1} - \varphi),$$

c'est-à-dire $z_n$ se déduit de l'équation

$$\Delta z = Lx_{n-1} - \varphi, \quad z|_S = 0.$$

D'après (3), on trouve pour $\varepsilon_n$

$$\varepsilon_n = \frac{(z_n, z_n)}{(Uz_n, z_n)} = \frac{\displaystyle\iint_D \left[ \left( \frac{\partial z_n}{\partial s} \right)^2 + \left( \frac{\partial z_n}{\partial t} \right)^2 \right] ds\, dt}{\displaystyle\iint_D \left[ a \left( \frac{\partial z_n}{\partial s} \right)^2 + b \left( \frac{\partial z_n}{\partial t} \right)^2 + c z_n^2 \right] ds\, dt}.$$

Pour la convergence du processus, le théorème 1.2, compte tenu de (7), donne

$$\| x_n - x^* \| \leqslant C \left( \frac{M-m}{M+m} \right)^n \leqslant C \left( \frac{\beta - \alpha}{\beta + \alpha} \right)^n \quad (n = 0, 1, \ldots), \tag{8}$$

où $x^*$ est solution de l'équation (2) ou, ce qui revient au même, du problème aux limites (1).

**3.2.** Le processus des approximations successives est d'une grande efficacité si les coefficients de l'équation donnée, les fonctions $a$, $b$. $c$, et le second membre $\varphi$ sont des polynômes. Dans ce cas, en prenant un polynôme pour $x_0$, on obtient pour la détermination de $z_1$ l'équation

$$\Delta z = p, \tag{9}$$

où $p$ est un polynôme. Il est immédiat de voir que la solution de cette équation (qui appartient à $\overset{\circ}{W}{}_2^{(1)}$) est également un polynôme dont le degré est de deux supérieur à celui du polynôme $p$. En effet, mettons $z$ sous la forme

$$z(s, t) = (s^2 + t^2 - 1)\,\pi(s, t), \tag{10}$$

où $\pi$ est un polynôme de coefficients indéterminés de même degré que $p$. En portant l'expression (10) dans l'équation (9) et en identifiant les coefficients du premier et du second membre, on obtient pour la détermination des coefficients du polynôme $\pi$ un système d'équations linéaires algébriques à un nombre d'inconnues égal à celui des équations. La matrice du système indiqué ne dépend visiblement pas de $p$. La solution polynomiale de l'équation (9) étant unique si elle existe, le déterminant du système est différent de zéro, donc l'équation (9) admet une solution polynomiale quel que soit le polynôme $p$ *).

Ainsi la fonction $z_1$ est un polynôme, donc $x_1 = x_0 - \varepsilon_1 z$ l'est, et, par suite, toutes les approximations $x_n$.

Cette circonstance permet, sous réserve que l'équation (1) soit correcte, d'établir aussi bien l'existence de la solution généralisée, que de la solution au sens classique du terme. Pour cela nous aurons besoin de certains résultats de la théorie des fonctions.

**3.3.** Indiquons tout d'abord l'analogue bidimensionnel des inégalités de Bernstein et de A. Markov. Soit $p$ un polynôme de degré $n$, tel que

$$|p(s, t)| \leqslant M \quad (s, t \in D),$$

*) En fait, pour résoudre l'équation (9), il est plus facile de chercher d'abord la solution sans se soucier des conditions aux limites, ensuite d'en soustraire la solution de l'équation de Laplace vérifiant les mêmes conditions pour trouver la solution vérifiant les conditions aux limites nulles. Cette dernière est particulièrement facile à trouver grâce à un développement en série de Fourier.

où $D$ est un domaine borné quelconque de contour différentiable. Alors dans $D$

$$\left|\frac{\partial p(s, t)}{\partial s}\right| \leqslant An^2M, \quad \left|\frac{\partial p(s, t)}{\partial t}\right| \leqslant An^2M \tag{11}$$

et dans tout domaine $D_1 \subset D$

$$\left|\frac{\partial p(s, t)}{\partial s}\right| \leqslant BnM, \quad \left|\frac{\partial p(s, t)}{\partial t}\right| \leqslant BnM, \tag{12}$$

où $A$ et $B$ sont des constantes dépendant uniquement de $D$ et $D_1$.

Les inégalités (11) et (12) découlent manifestement des inégalités correspondantes en dimension un (cf. Natanson [I], pages 169 et 174).

Prouvons le lemme suivant.

LEMME 1. *Si $u$ est un polynôme de degré $\leqslant n$,*

$$\| u \|_{C(D)} \leqslant A_1 n^2 \| u \|_{L^2(D)}, \quad \| u \|_{C(D_1)} \leqslant B_1 n \| u \|_{L^2(D)}, \tag{13}$$

$$\| u \|_{C^{(1)}(D)} \leqslant A_2 n^2 \| u \|_{W_2^{(1)}(D)}, \quad \| u \|_{C^{(1)}(D_1)} \leqslant B_1 n \| u \|_{W_2^{(1)}(D)}, \tag{14}$$

*où $D_1 \subset D$ et $C^{(1)}$ est l'espace des fonctions continûment différentiables.*

DÉMONSTRATION. Limitons-nous au cas où le domaine $D$ est un disque centré en l'origine des coordonnées (le cas général implique quelques raisonnements supplémentaires). En posant

$$\varphi(s, t) = \int_0^s \int_0^t p^2(s, t)\, ds\, dt,$$

on obtient

$$|\varphi(s, t)| \leqslant \| p \|_{L^2}^2 \quad ((s, t) \in D),$$

et après une double application de l'inégalité (11)

$$|p(s, t)|^2 = \left|\frac{\partial^2 \varphi(s, t)}{\partial s\, \partial t}\right| \leqslant$$
$$\leqslant A^2 (2n+2)^2 (2n+1)^2 \max |\varphi(s, t)| \leqslant A_1^2 n^4 \| p \|_{L^2}^2,$$

d'où l'on déduit la première des inégalités (13). La deuxième inégalité s'établit de façon analogue sauf qu'il faut se servir de (12) au lieu de (11).

En appliquant l'inégalité établie à la dérivée, on trouve

$$\left|\frac{\partial p(s, t)}{\partial s}\right| \leqslant A_1 n^2 \left\| \frac{\partial p}{\partial s} \right\|_{L^2} \leqslant A_1 n^2 \| p \|_{W_2^{(1)}};$$

la même inégalité est valable pour la dérivée $\frac{\partial p}{\partial t}$. Enfin, en vertu de (11) et du théorème d'immersion,

$$|p(s, t)| \leqslant An^2 \| p \|_{L^2} \leqslant CAn^2 \| p \|_{W_2^{(1)}}.$$

En combinant tout ceci on est conduit à la première des inégalités (14):

$$\| p \|_{C^{(1)}(D)} = \max_D | p(s, t) | + \max_D \left| \frac{\partial p(s, t)}{\partial s} \right| + \max_D \left| \frac{\partial p(s, t)}{\partial t} \right| \leqslant$$

$$\leqslant A_2 n^2 \| p \|_{W_2^{(1)}}.$$

La deuxième inégalité s'établit de façon analogue.

**3.4.** Revenons à l'équation (1). On admettra que les coefficients $a$, $b$, $c$ et le second membre $\varphi$ sont des polynômes. Comme indiqué dans 3.2, si l'approximation initiale $x_0$ est un polynôme, toutes les approximations suivantes $x_n$ ($n = 1, 2, \ldots$) le seront aussi. De plus, si les polynômes $a$, $b$, $c$ et $\varphi$ sont de degré $\leqslant m$, et $x_0$ un polynôme de degré $N$, alors $Lx_0$ sera de degré $\leqslant m + N$, et $z_1$ de degré $\leqslant m + N + 2$. En particulier, si $x_0 = 0$, alors $z_1$ et, par suite, $x_1$ seront de degré $\leqslant m + 2$. En poursuivant ces raisonnements, on trouve que $x_n$ est un polynôme de degré $\leqslant (m + 2) n$.

En tenant compte de la majoration (8), on obtient

$$\| x_k - x_{k-1} \|_{\overset{\circ}{W}{}_2^{(1)}} \leqslant \| x_k - x^* \| + \| x_{k-1} - x^* \| \leqslant$$

$$\leqslant 2C \left( \frac{\beta - \alpha}{\beta + \alpha} \right)^{k-1} \quad (k = 1, 2, \ldots).$$

La différence $x_k - x_{k-1}$ étant un polynôme de degré $\leqslant (m + 2) k$, le lemme nous dit que

$$\| x_k - x_{k-1} \|_{C^{(1)}(D)} \leqslant A_3 (m+2)^2 k^2 q \quad (k = 1, 2, \ldots),$$

où

$$q = \frac{\beta - \alpha}{\beta + \alpha}.$$

En appliquant l'inégalité (11) plusieurs fois, on obtient la majoration

$$\| x_k - x_{k-1} \|_{C^{(p)}(D)} \leqslant A_4 [(m+2) k]^{2p} q^k \quad (k, p = 1, 2, \ldots). \quad (15)$$

Donc, la série

$$x^*(s, t) = \sum_{k=1}^{\infty} [x_k(s, t) - x_{k-1}(s, t)] \quad (16)$$

peut être dérivée $p$ fois terme à terme. Il est équivalent de dire que la solution de l'équation (1) admet des dérivées de tout ordre et de plus

$$\| x^* - x_n \|_{C^{(p)}} = \sum_{k=n+1}^{\infty} \| x_k - x_{k-1} \| \leqslant A_5 [(m+2) n]^{2p} q^n \quad (17)$$

$$(n = 0, 1, \ldots),$$

c'est-à-dire la suite $\{x_n\}$ des approximations obtenues par la méthode de la plus forte pente converge vers la solution $x^*$ uniformément avec toutes ses dérivées.

D'après ce qui précède il y a lieu de noter que si l'on utilise le théorème de Bernstein relatif aux approximations d'une fonction analytique par des polynômes (Natanson [I], page 228), on peut établir l'analyticité de la solution $x^*$.

Passons maintenant au cas où les coefficients de l'équation (1) et son second membre sont des fonctions plusieurs fois dérivables. Plus exactement, on admettra que les fonctions $a$ et $b$ possèdent des dérivées jusqu'à l'ordre $(\nu + 1)$, les fonctions $c$ et $\varphi$, jusqu'à l'ordre $\nu$, et de plus que les dérivées d'ordre $\nu + 1$ des fonctions $a$ et $b$ et les dérivées d'ordre $\nu$ des fonctions $c$ et $\varphi$ vérifient la condition de Lipschitz dans le rapport $\alpha > 0$. D'après le théorème de Jackson (cf. par exemple Harrik [1]) on peut trouver des polynômes $a_n$, $b_n$, $c_n$ et $\varphi_n$ de degré $\leqslant n$ tels que

$$|a(s, t) - a_n(s, t)| \leqslant \frac{K}{n^{\nu+1+\alpha}},$$

$$\left|\frac{\partial}{\partial s}[a(s, t) - a_n(s, t)]\right| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$\left|\frac{\partial}{\partial t}[a(s, t) - a_n(s, t)]\right| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$|b(s, t) - b_n(s, t)| \leqslant \frac{K}{n^{\nu+1+\alpha}},$$

$$\left|\frac{\partial}{\partial s}[b(s, t) - b_n(s, t)]\right| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$\left|\frac{\partial}{\partial t}[b(s, t) - b_n(s, t)]\right| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$|c(s, t) - c_n(s, t)| \leqslant \frac{K}{n^{\nu+\alpha}},$$

$$|\varphi(s, t) - \varphi_n(s, t)| \leqslant \frac{K}{n^{\nu+\alpha}}.$$

Désignons par $L_n$ l'opérateur déduit de $L$ par substitution de $a_n$, $b_n$, $c_n$ à $a$, $b$, $c$ respectivement. Il est clair que si $n$ est assez grand, on peut appliquer la méthode de la plus forte pente à l'opérateur $L_n$ et admettre de plus que $\alpha$ et $\beta$ restent valables pour l'opérateur $L_n$. Désignons par $x^{(n)}$ une solution de l'équation

$$L_n x = \varphi_n, \quad x|_S = 0, \tag{18}$$

et évaluons la différence de deux telles solutions

$$\| Lx^{(n)} - L_n x^{(n)} \|_{L^2} \leqslant$$

$$\leqslant \left\| \frac{\partial}{\partial s} (a - a_n) \frac{\partial x^{(n)}}{\partial s} \right\| + \left\| (a - a_n) \frac{\partial^2 x^n}{\partial s^2} \right\| +$$

$$+ \left\| \frac{\partial}{\partial t} (b - b_n) \frac{\partial x^n}{\partial t} \right\| + \left\| (b - b_n) \frac{\partial^2 x^{(n)}}{\partial t^2} \right\| + \| (c - c_n) x^{(n)} \| \leqslant$$

$$\leqslant \frac{K_1}{n^{\nu+\alpha}} \left[ \left\| \frac{\partial x^{(n)}}{\partial s} \right\|_{L^2} + \left\| \frac{\partial^2 x^{(n)}}{\partial s^2} \right\|_{L^2} + \left\| \frac{\partial x^{(n)}}{\partial t} \right\|_{L^2} + \right.$$

$$\left. + \left\| \frac{\partial^2 x^{(n)}}{\partial t^2} \right\|_{L^2} + \| x^{(n)} \|_{L^2} \right] \leqslant \frac{K_2}{n^{\nu+\alpha}} \| x^{(n)} \|_{\overset{\circ}{W}{}_2^{(2)}}.$$

On a ensuite

$$\| L (x^{(n)} - x^{(m)}) \|_{L^2} \leqslant$$

$$\leqslant \| (L - L_n) x^{(n)} \| + \| (L - L_m) x^{(m)} \| + \| \varphi_n - \varphi_m \| \leqslant$$

$$\leqslant \frac{K_2}{n^{\nu+\alpha}} \| x^{(n)} \|_{\overset{\circ}{W}{}_2^{(2)}} + \frac{K_2}{m^{\nu+\alpha}} \| x^{(m)} \|_{\overset{\circ}{W}{}_2^{(2)}} + \frac{K}{n^{\nu+\alpha}} + \frac{K}{m^{\nu+\alpha}}.$$

D'où, en vertu de l'inégalité de Bernstein-Ladyjenskaïa (cf. Ladyjenskaïa, chap. II, § 6), l'on obtient

$$\| x^{(n)} - x^{(m)} \|_{\overset{\circ}{W}{}_2^{(2)}} \leqslant K_3 \| L (x^{(n)} - x^{(m)}) \|_{L^2} \leqslant$$

$$\leqslant K_4 \left[ \frac{1}{n^{\nu+\alpha}} \| x^{(n)} \|_{\overset{\circ}{W}{}_2^{(2)}} + \frac{1}{m^{\nu+\alpha}} \| x^{(m)} \|_{\overset{\circ}{W}{}_2^{(2)}} + \frac{1}{n^{\nu+\alpha}} + \frac{1}{m^{\nu+\alpha}} \right]. \quad (19)$$

En appliquant cette majoration au couple $x^{(n_0)}$, $x^{(m)}$ ($n_0$ est un nombre fixe assez grand) et en remplaçant à droite $\| x^{(m)} \|$ par la somme plus grande $\| x^{(n_0)} \| + \| x^{(n_0)} - x^{(m)} \|$, on établit que la quantité $\| x^{(n_0)} - x^{(m)} \|_{\overset{\circ}{W}{}_2^{(2)}}$ est bornée indépendamment de $m$, donc $\| x^{(m)} \|$ l'est aussi. En supposant que $n < m$ dans (19) et en utilisant le fait que $\|x^{(m)}\|$ est bornée, mettons (19) sous la forme

$$\| x^{(n)} - x^{(m)} \|_{\overset{\circ}{W}{}_2^{(2)}} \leqslant \frac{K_5}{n^{\nu+\alpha}}. \quad (20)$$

Prenons un nombre $h > 1$ (plus bas nous nous étendrons sur le choix de ce nombre) et considérons la suite $\{n_k\}$ des entiers naturels $n_k = [h^k]$ ($[\gamma]$ est la partie entière de $\gamma$, $k = 1, 2, \ldots$). Construisons pour chaque fonction $x^{(n_k)}$ la $k$-ième approximation $x_k^{(n_k)}$ d'après la méthode de la plus forte pente (les approximations initiales sont nulles). Comme indiqué plus haut, $x_k^{(n_k)}$ est un polynôme de degré $\leqslant (n_k + 2) k$. Evaluons l'écart de deux tels polynômes. D'abord

$$\| x_k^{(n_k)} - x_{k-1}^{(n_{k-1})} \| \leqslant \| x_k^{(n_k)} - x^{(n_k)} \| +$$

$$+ \| x_{k-1}^{(n_{k-1})} - x^{(n_{k-1})} \| + \| x^{(n_k)} - x^{(n_{k-1})} \|.$$

De (20) il résulte que

$$\| x^{(n_k)} - x^{(n_{k-1})} \|_{\overset{\circ}{W}{}_2^{(2)}} \leqslant \frac{K_5}{n_{k-1}^{\nu+\alpha}} \leqslant \frac{K_6}{n_k^{\nu+\alpha}}. \tag{21}$$

Ensuite, d'après (17)

$$\| x^{(n_k)} - x_k^{(n_k)} \|_{C^{(2)}} \leqslant A_5 [(n_k + 2) k]^4 q^k \quad (k = 1, 2, \ldots). \tag{22}$$

Donc

$$\| x_k^{(n_k)} - x_{k-1}^{(n_{k-1})} \|_{\overset{\circ}{W}{}_2^{(2)}} \leqslant K_7 [(n_k + 2) k]^4 q^k + \frac{K_6}{n_k^{\nu+\alpha}}$$

Comme sous le signe de la norme figure un polynôme de degré $\leqslant (n_k + 2) k$, le lemme et l'inégalité (11) nous donnent

$$\| x_k^{(n_k)} - x_{k-1}^{(n_{k-1})} \|_{C^{(p)}} \leqslant$$
$$\leqslant K_8 [(n_k + 2) k]^{2p-2} \left\{ K_7 [(n_k + 2) k]^4 q^k + \frac{K_6}{n_k^{\nu+\alpha}} \right\} = \varepsilon_k.$$

Comme $n_k \leqslant h^k$, il est clair que si $\nu + \alpha > 2p - 2$ et que $h$ soit choisi tel que $h^{2p+2} q < 1$, alors la série $\sum_{k=1}^{\infty} \varepsilon_k$ est convergente, donc la suite $\{x_k^{(n_k)}\}$ l'est également dans l'espace $C^{(p)}$ et sa limite est un élément de $C^{(p)}$, c'est-à-dire une fonction $p$-continûment différentiable

$$x(s, t) = \lim_{k \to \infty} x_k^{(n_k)}(s, t).$$

Si $p \geqslant 2$, en utilisant les mêmes majorations, on vérifie immédiatement que $x$ est solution de l'équation (1), c'est-à-dire est solution classique du problème aux limites. Donc, la solution classique existe si $\nu + \alpha > 2$ et notamment pour $\nu = 2$, c'est-à-dire si les coefficients $a$ et $b$ et si $c$ et $\varphi$ admettent des dérivées troisièmes et secondes respectivement vérifiant la condition de Lipschitz dans un rapport $\alpha > 0$.

D'une manière analogue, en utilisant les majorations pour les sous-domaines intérieurs, on trouve que la solution est 2-continûment différentiable à l'intérieur du domaine déjà pour $\nu = 1$.

A noter que cette méthode peut être appliquée à l'étude de problèmes plus compliqués. Le cas d'un domaine non circulaire se ramène au cas étudié si ce domaine s'applique continûment sur un disque. Moyennant des changements insignifiants, la méthode peut être utilisée pour des équations d'ordre supérieur, à un nombre de variables indépendantes plus grand que deux.

## § 4. Minimisation des fonctionnelles convexes différentiables

**4.1.** On se propose d'appliquer la méthode de la plus forte pente à la minimisation de fonctionnelles différentiables. Une fonctionnelle $\Phi$ définie sur un $B$-espace X est différentiable en un point $x \in \mathbf{X}$ si existe une fonctionnelle linéaire $f$ telle que pour $h \in \mathbf{X}$ l'on ait

$$\Phi(x+h) = \Phi(x) + f(h) + o(h),$$

où $o(h)/\|h\| \to 0$ pour $\|h\| \to 0$. La fonctionnelle $f$ est appelée *dérivée* de Fréchet de la fonctionnelle $\Phi$ en $x$ et se note $\Phi'(x)$. (Pour plus de détails sur la différentiation, cf. chap. XVII.)

Soit X un $B$-espace. Dans les problèmes d'extrémums il y a intérêt à considérer (contrairement aux paragraphes précédents) que l'espace X est réel. Considérons une fonctionnelle $\Phi$ définie et différentiable sur X et déterminons la direction de plus forte pente en un point $x \in \mathbf{X}$. La fonctionnelle $\Phi$ étant différentiable, sa dérivée au point $x$ dans la direction de $z$ s'écrit

$$\frac{\partial\Phi}{\partial z}(x) = \frac{1}{\|z\|} \lim_{\alpha\to 0+} \frac{1}{\alpha}(\Phi(x+\alpha z) - \Phi(x)) =$$
$$= \frac{1}{\|z\|}\Phi'(x)(z) = \Phi'(x)\left(\frac{z}{\|z\|}\right).$$

De ce qui précède il suit que le vecteur unité $y$ de la plus forte pente de la fonctionnelle $\Phi$ au point $x$ est confondu avec le point minimum de la fonctionnelle linéaire $\Phi'(x)$ sur la sphère unité, c'est-à-dire se détermine à partir de la condition

$$\Phi'(x)(y) = \inf_{\|z\|=1} \Phi'(x)(z). \tag{1}$$

La fonctionnelle linéaire n'atteignant pas nécessairement son minimum sur la sphère unité, il peut ne pas exister de direction de plus forte pente. Toutefois, il n'est pas difficile d'indiquer les conditions sous lesquelles le minimum est réalisé dans (1): par exemple, si X est un espace dual d'un $B$-espace Y, et les fonctionnelles $\Phi'(x)$ ($x \in \mathbf{X}$) appartiennent à Y. (Plus exactement, à l'image de Y par l'injection canonique de Y dans $\mathbf{Y}^{**} = \mathbf{X}^*$). En effet, la boule unité dans $\mathbf{X} = \mathbf{Y}^*$ étant compacte pour la topologie $(*)$-faible, $\Phi'(x)$ atteint son minimum sur cette boule, donc sur la sphère. En particulier, si X est réflexif, la direction de plus forte pente existe manifestement.

On conviendra dans la suite sans le mentionner expressément que X et $\Phi$ sont tels que la direction de plus forte pente existe.

Si l'espace X est strictement convexe *), la relation (1) nous dit que la direction de plus forte pente est unique. Ceci bien entendu n'est pas valable dans le cas général.

La formule (1) montre que la direction de plus forte pente (et par conséquent la réalisation de la méthode) dépend essentiellement et de la fonctionnelle et de la norme de l'espace. Le passage à une autre norme, même équivalente à l'originelle, peut grandement modifier cette direction. La formule (1) et l'égalité

$$\inf_{\| z \| = 1} \Phi'(x)(z) = - \sup_{\| z \| = 1} (-\Phi'(x)(z)) = -\| \Phi'(x) \|$$

entraînent que la direction de $y$ ($\| y \| = 1$) est direction de plus forte pente si, et seulement si,

$$\Phi'(x)(y) = -\| \Phi'(x) \|. \tag{2}$$

Pour tout $z \in X$ on a

$$-\Phi'(x)(z) \leqslant \| \Phi'(x) \| \, \| z \|.$$

La direction de plus forte pente est caractérisée par le fait qu'elle réalise l'égalité dans la relation précédente. Ceci permet dans certains cas de trouver assez facilement cette direction. Exhibons quelques exemples.

1) Soit X un espace hilbertien. L'égalité $-\Phi'(x)(z) = \| \Phi'(x) \| \, \| z \|$. ou ce qui revient au même $(-\Phi'(x), z) = \| -\Phi'(x) \| \, \| z \|$, montre que l'égalité est réalisée dans l'inégalité de Cauchy-Bouniakovski pour les vecteurs $-\Phi'(x)$ et $z$. Ce qui exprime que ces vecteurs sont positivement proportionnels: il existe $\lambda > 0$ tel que $-\Phi'(x) = \lambda z$. Donc, la direction de plus forte pente est définie par le vecteur $-\Phi'(x)$.

Remarque. Si X est un espace hilbertien, la dérivée $\Phi'(x)$ peut être considérée comme un de ses éléments. Cet élément est généralement appelé *gradient* de $\Phi$ au point $x$. Dans le même ordre d'idées la direction de plus forte pente est appelée direction de l'*antigradient*.

2) Soit $X = L^p(a, b)$ $(1 < p < \infty)$. Une analyse de la condition de réalisabilité de l'égalité dans l'inégalité de Hölder nous permet de vérifier immédiatement que la direction de plus forte pente de $z$ en $x$ est donnée par la formule

$$z(t) = -[\operatorname{sign} y(t)] |y(t)|^{q-1},$$

---

*) Un $B$-espace X est *strictement convexe* si dans l'inégalité triangulaire l'égalité est réalisée seulement pour des éléments positivement proportionnels, c'est-à-dire la relation $\| x + y \| = \| x \| + \| y \|$ entraîne que $x = \lambda y$ pour un $\lambda > 0$. Si X est strictement convexe, $x, y \in X$ et $\| x \| = \| y \| = \frac{1}{2} \| x + y \| = 1$, alors $x = y$. L'espace $L^p(a, b)$ $(1 < p < +\infty)$ est un espace strictement convexe.

où $1/p + 1/q = 1$ et $y$ est un élément de l'espace $\mathbf{L}^q(a, b)$ tel que

$$\Phi'(x)(u) = \int_a^b y(t)\,u(t)\,dt, \quad \forall u \in \mathbf{L}^p(a, b).$$

Comme dans le cas précédent, il existe une seule direction de plus forte pente.

3) Supposons que $\mathbf{X} = \mathbf{L}^\infty(a, b)$ et que les fonctionnelles $\Phi'(x)$ $(x \in \mathbf{L}^\infty(a, b))$ appartiennent à $\mathbf{L}^1(a, b)$ (c'est-à-dire $\Phi'(x)(u) = \int_a^b y(t)\,u(t)\,dt$, où $y \in \mathbf{L}^1(a, b)$). Il existe généralement plusieurs directions de pente maximum dans ce cas. Chacune d'elles est donnée par une fonction mesurable $z$ telle que vrai sup $|z(t)| = 1$ et $z(t) = -\text{sign}\, y(t)$ si $y(t) \neq 0$. (Si l'ensemble $\{t : y(t) = 0\}$ est de mesure nulle, cette direction est unique.)

4.2. Indiquons quelques définitions avant de passer à l'étude de la méthode de la plus forte pente.

Un point $x \in \mathbf{X}$ est *point de minimum local* d'une fonctionnelle $\Phi$ si $\Phi(z) \geqslant \Phi(x)$ pour les éléments $z$ d'une boule centrée en $x$. Si $\Phi(z) \geqslant \Phi(x)$ pour tous les $z \in \mathbf{X}$, on dit que $x$ est *point de minimum global*. Un point $x \in \mathbf{X}$ est un *point stationnaire* d'une fonctionnelle dérivable $\Phi$ si $\Phi'(x) = 0$.

Les fonctions d'une variable réelle $\varphi(\alpha; x, z) = \Phi(x + \alpha z)$ envisagées dans 1.1 jouent un rôle important dans la recherche du minimum.

Il est immédiat de vérifier que la dérivabilité de la fonctionnelle $\Phi$ entraîne celle (par rapport à $\alpha$) de la fonction $\varphi(\alpha; x, z)$, quels que soient $x$ et $z$, et de plus $\varphi'(\alpha; x, z) = \Phi'(x + \alpha z)(z)$. Utilisons ceci pour prouver que le minimum local est un point stationnaire. Si $x$ est un point de minimum local, alors pour tout $z \in \mathbf{X}$ la fonction $\varphi(\alpha; x, z)$ atteint un minimum local sur la droite réelle en 0; donc, $\varphi'(0; x, z) = \Phi'(x)(z) = 0$. Or, cela exprime que le point $x$ est stationnaire.

La réciproque est manifestement fausse. Cependant, si $\Phi$ est une fonctionnelle convexe, les points stationnaires sont confondus avec les points de minimum global.

Une fonctionnelle $\Phi$ définie sur un espace vectoriel $\mathbf{X}$ est *convexe* si pour tous éléments $z_1, z_2 \in \mathbf{X}$ et tous nombres $t_1, t_2 \geqslant 0$ tels que $t_1 + t_2 = 1$, on a

$$\Phi(t_1 x_1 + t_2 x_2) \leqslant t_1 \Phi(x_1) + t_2 \Phi(x_2).$$

Si $\Phi$ est une fonctionnelle convexe, la fonction $\varphi(\alpha; x, z)$ est convexe sur la droite réelle pour tous $x$ et $z$, c'est-à-dire

$$\varphi(t_1\alpha_1 + t_2\alpha_2; x, z) \leqslant t_1\varphi(\alpha_1; x, z) + t_2\varphi(\alpha_2; x, z),$$

où $t_1$, $t_2$, $\alpha_1$, $\alpha_2$ sont des nombres réels; $t_1, t_2 \geqslant 0$, $t_1 + t_2 = 1$. La convexité de $\varphi(\alpha; x, z)$ entraîne la croissance de la dérivée de cette fonction.

LEMME 1. *Soit $\Phi$ une fonctionnelle différentiable convexe. Alors chaque point stationnaire de $\Phi$ est point de minimum global.*

DÉMONSTRATION. Soient $x$ un point stationnaire, $z$ un élément quelconque de X. La dérivée de la fonction $\varphi(\alpha; x, z)$ étant croissante, la formule de Lagrange nous dit que pour un $\theta \in ]0, 1[$

$$\Phi(x+z) - \Phi(z) = \varphi(1; x, z) - \varphi(0; x, z) = \\ = \varphi'(\theta, x, z) \geqslant \varphi'(0; x, z) = \Phi'(x)(z) = 0.$$

Donc, $\Phi(x+z) \geqslant \Phi(x)$, $\forall z \in X$, ce qui prouve le lemme.

Dans la suite, nous étudions la convergence de la méthode de la plus forte pente vers les points stationnaires de la fonctionnelle $\Phi$ (donc vers les points de minimum global si $\Phi$ est convexe).

Les principaux résultats de convergence ne subsistent que si la dérivée de la fonctionnelle $\Phi$ vérifie la condition de Lipschitz dans une boule, c'est-à-dire existe un nombre $L$ tel que pour $x$ et $z$ appartenant à cette boule l'on ait

$$\| \Phi'(x) - \Phi'(z) \| \leqslant L \| x - z \|. \tag{3}$$

On se servira du lemme suivant.

LEMME 2. *Supposons que la dérivée $\Phi'$ vérifie la condition de Lipschitz* (3) *dans la boule $B_{a+b}$* (0) *de rayon $a + b$ et de centre* 0. *Si $\| x \| \leqslant a$, alors*

$$\Phi(x+z) \leqslant \Phi(z) + \Phi'(x)(z) + \frac{L}{2} \| z \|^2.$$

DÉMONSTRATION. En utilisant la dérivabilité de la fonction $\varphi(\alpha; x, z)$ et l'expression de sa dérivée, on obtient

$$\Phi(x+z) = \varphi(1; x, z) = \varphi(0; x, z) + \int_0^1 \varphi'(\alpha; x, z)\, d\alpha =$$

$$= \Phi(x) + \int_0^1 \Phi'(x+\alpha z)(z)\, d\alpha =$$

$$= \Phi(x) + \Phi'(x)(z) + \int_0^1 (\Phi'(x+\alpha z) - \Phi'(x))(z)\, d\alpha \leqslant$$

$$\leqslant \Phi(x) + \Phi'(x)(z) + \int_0^1 \| \Phi'(x+\alpha z) - \Phi'(x) \| \, \| z \| \, d\alpha \leqslant$$

$$\leqslant \Phi(x) + \Phi'(x)(z) + L \| z \|^2 \int_0^1 \alpha \, d\alpha = \Phi(x) + \Phi'(x)(z) + (L/2) \| z \|^2,$$

c.q.f.d.

**4.3.** Utilisons la méthode de la plus forte pente pour minimiser la fonctionnelle $\Phi$, ou plus exactement étudions le comportement de la suite $x_0, x_1, \ldots, x_n, \ldots$, où

$$x_n = x_{n-1} + \varepsilon_n z_n \quad (n = 1, 2, \ldots), \tag{4}$$

$z_n$ étant la direction de plus forte pente au point $x_{n-1}$ (s'il y a plusieurs telles directions, $z_n$ sera l'une d'elles), $\varepsilon_n$ se déterminant à partir de la condition

$$\Phi(x_{n-1} + \varepsilon_n z_n) = \min_{\alpha \geqslant 0} \Phi(x_{n-1} + \alpha z_n).$$

Donc, la valeur de la pente est celle donnée par la deuxième méthode de 1.1. On remarquera par ailleurs que si $\Phi$ est une fonctionnelle strictement convexe, les deux premières méthodes de 1.1 sont confondues. (La convexité stricte de $\Phi$ signifie que $\Phi(t_1 x_1 + t_2 x_2) < t_1 \Phi(x_1) + t_2 \Phi(x_2)$ pour $x_1, x_2 \in X$, $x_1 \neq x_2$; $t_1, t_2 > 0$, $t_1 + t_2 = 1$; la convexité stricte de $\Phi$ entraîne que la dérivée $\varphi'(\alpha; x, z)$ est strictement croissante et ne peut donc s'annuler qu'en un seul point; d'où l'équivalence des méthodes mentionnées.)

On admet dans la suite sans mention spéciale que le point initial $x_0 \in X$ est tel que l'ensemble de Lebesgue $\Omega_0 = \{x \in X : \Phi(x) \leqslant \Phi(x_0)\}$ est borné. On remarquera que $x_n \in \Omega_0$ $(n = 0, 1, 2, \ldots)$. On conviendra également que $\| z_n \| = 1$ pour tous les $n$.

THÉORÈME 1. *Supposons que la dérivée de la fonctionnelle $\Phi$ vérifie la condition de Lipschitz* (3) *dans une boule de centre* 0 *et de rayon $R'$, où $R' > R = \sup_{x \in \Omega_0} \| x \|$. Alors $\Phi'(x_n) \to 0$, où $\{x_n\}$ est une suite construite à l'aide de la formule* (4).

DÉMONSTRATION. Soit $0 < \alpha \leqslant R' - R$. Alors par définition de la valeur de la pente $\varepsilon_n$ on a $\Phi(x_{n-1} + \alpha z_n) \geqslant \Phi(x_n)$. Le lemme 2 nous dit que

$$\Phi(x_n) \leqslant \Phi(x_{n-1} + \alpha z_n) \leqslant \Phi(x_{n-1}) - \alpha \Phi'(x_{n-1})(z_n) + \frac{1}{2} L \alpha^2$$

(où $L$ est le rapport de la condition de Lipschitz). Ceci et l'inégalité (2) entraînent que

$$\| \Phi'(x_n) \| \leqslant \frac{\Phi(x_{n-1}) - \Phi(x_n)}{\alpha} + \frac{1}{2} L \alpha.$$

Soient $\varepsilon$ un nombre strictement positif et $\alpha < \min(\varepsilon, R' - R)$. Le fait que $\Omega_0$ soit bornée et la condition de Lipschitz impliquent que la fonctionnelle $\Phi$ est minorée. La suite $\Phi(x_n)$ étant décroissante (par construction) et bornée, elle est convergente, donc pour les $n$ assez grands on a $\frac{1}{\alpha}(\Phi(x_{n-1}) - \Phi(x_n)) < \varepsilon$. Pour de tels $n$ on a

$$\|\Phi'(x_n)\| \leqslant \varepsilon\left(1 + \frac{1}{2}L\right),$$

c.q.f.d.

COROLLAIRE. *Les points limites de la suite* (4) (*sous réserve qu'ils existent*) *sont stationnaires.*

Dans certains cas, en se servant de la compacité, on peut prouver la convergence vers un point stationnaire pour des fonctionnelles dont la dérivée n'est que continue. Indiquons un résultat de cette nature (Curry [1]).

THÉORÈME 2. *Supposons que* X *est de dimension finie et la fonctionnelle* $\Phi$ *continûment dérivable. Alors les points limites de la suite* $\{x_n\}$ *sont stationnaires.*

DÉMONSTRATION. Signalons tout d'abord que la suite $\{x_n\}$ admet des points limites, car X est de dimension finie et $\Omega_0$ est borné. Soit $y = \lim x_{n_k}$ l'un de ces points. Sans restreindre la généralité nous admettons qu'existe $\lim z_{n_k+1} = \tilde{y}$. Posons pour les $\alpha > 0$

$$w_{n_k}(\alpha) = \frac{1}{\alpha}(\Phi(x_{n_k} + \alpha z_{n_k+1}) - \Phi(x_{n_k})) - \Phi'(x_{n_k})(z_{n_k+1}). \quad (5)$$

Par hypothèse, il existe $\lim\limits_{k\to\infty} w_{n_k}(\alpha) = w(\alpha)$, et de plus

$$w(\alpha) = \frac{1}{\alpha}(\Phi(y + \alpha\tilde{y}) - \Phi(y)) - \Phi'(y)(\tilde{y}).$$

$\Phi$ étant dérivable au point $y$, on a $\lim\limits_{\alpha\to 0} w(\alpha) = 0$. D'après le choix de la valeur de la pente,

$$\Phi(x_{n_k+1}) = \Phi(x_{n_k} + \varepsilon_{n_k+1} z_{n_k+1}) \leqslant \Phi(x_{n_k} + \alpha z_{n_k+1}),$$

donc, en se servant de (5) on obtient

$$\Phi(x_{n_k+1}) \leqslant \Phi(x_{n_k}) - \alpha\Phi'(x_{n_k})(z_{n_k+1}) + \alpha w_{n_k}(\alpha).$$

Comme $\alpha > 0$, il vient

$$-\Phi'(x_{n_k})(z_{n_k+1}) \leqslant \frac{1}{\alpha}(\Phi(x_{n_k}) - \Phi(x_{n_k+1})) + w_{n_k}(\alpha). \quad (6)$$

Dans (6) passons à la limite pour $k \to \infty$. En tenant compte de ce que $\lim(\Phi(x_{n_k}) - \Phi(x_{n_k+1})) = 0$ et en utilisant (2), on obtient

$$\|\Phi'(y)\| = -\Phi'(y)\tilde{y} \leqslant w(\alpha).$$

La proposition du théorème découle maintenant de la relation $w(\alpha) \underset{\alpha \to 0}{\longrightarrow} 0$.

4.4. Si la fonctionnelle $\Phi$ est convexe, on peut établir bien plus de résultats sur la convergence de la méthode de la plus forte pente.

LEMME 3. *Si $\Phi$ est une fonctionnelle convexe, minorée sur* X *et* $Q = \inf\limits_{x \in X} \Phi(x)$, *il existe un nombre $c > 0$ tel que*

$$\Phi(x_n) - Q \leqslant c \, \| \Phi'(x_n) \|.$$

DÉMONSTRATION. L'ensemble $\Omega_0 = \{z \in X : \Phi(z) \leqslant \Phi(x_0)\}$ étant borné par hypothèse, l'ensemble $\Omega_0 - \Omega_0 = \{z : z_1 = z_1 - z_2, z_1, z_2 \in \Omega_0\}$ l'est également. Montrons que pour $c$ on peut prendre le rayon d'une boule $B$ de centre 0 contenant $\Omega_0 - \Omega_0$. La dérivée $\varphi'(\alpha; x_n, z)$ étant décroissante (en vertu de la convexité de $\Phi$), la formule de Lagrange nous dit que

$$\Phi(x_n + z) - \Phi(x_n) = \varphi(1; x_n, z) - \varphi(0; x_n, z) =$$
$$= \varphi'(\theta; x_n, z) \geqslant \varphi'(0; x_n, z) = \Phi'(x_n)(z), \ \theta \in \, ]0, 1[.$$

D'où

$$\min_{\|z\| \leqslant c} \Phi(x_n + z) - \Phi(x_n) \geqslant \min_{\|z\| \leqslant c} \Phi'(x_n)(z). \tag{7}$$

Comme $x_n \in \Omega_0$, il vient $B \supset \Omega_0 - \Omega_0 \supset \Omega_0 - x_n$, c'est-à-dire $x_n + B \supset \Omega_0$. Donc

$$Q = \inf_{x \in X} \Phi(x) \leqslant \inf_{\|z\| \leqslant c} \Phi(x_n + z) \leqslant \inf_{x \in \Omega_0} \Phi(x).$$

De la définition de $\Omega_0$ il suit que $Q = \inf\limits_{x \in \Omega_0} \Phi(x)$. De plus

$$\inf_{\|z\| \leqslant c} \Phi'(x_n)(z) = c \inf_{\|z\| \leqslant 1} \Phi'(x_n)(z) = -c \, \| \Phi'(x_n) \|.$$

Les relations obtenues jointes à (7) entraînent que

$$Q - \Phi(x_n) \geqslant -c \, \| \Phi'(x_n) \|.$$

Ce qui prouve le lemme.

COROLLAIRE. *Si l'on se place dans les conditions du lemme* 3 *et de l'un des théorèmes* 1 *ou* 2, *la suite* $\{x_n\}$ *est minimisante.*

Si l'on admet que la dérivée $\Phi'(x)$ vérifie la condition de Lipschitz, on peut majorer la vitesse de convergence de la suite $\{\Phi(x_n)\}$.

THÉORÈME 3. *Si $\Phi$ est une fonctionnelle convexe, minorée sur* X, *dont la dérivée $\Phi'(x)$ vérifie sur* X *la condition de Lipschitz* (3), *alors* $\Phi(x_n) - Q = O(1/n)$, *où* $Q = \inf\limits_{x \in X} \Phi(x)$.

La démonstration s'appuie sur le lemme suivant.

LEMME 4. *Si les nombres* $\lambda_n > 0$ $(n = 1, 2, \ldots)$ *sont tels que pour un* $\mu > 0$ *et pour tous les* $n$ *l'on ait* $\lambda_n - \lambda_{n+1} \geqslant \mu\lambda_n^2$, *alors* $\lambda_n = O\left(\frac{1}{n}\right)$.

DÉMONSTRATION. La suite $\lambda_n$ étant décroissante, on a

$$\lambda_n - \lambda_{n+1} = \lambda_{n+1}\left(\frac{\lambda_n}{\lambda_{n+1}} - 1\right) \geqslant \mu\lambda_{n+1}^2,$$

d'où il suit que $\frac{\lambda_n}{\lambda_{n+1}} - 1 \geqslant \mu\lambda_{n+1}$.

En posant $\nu_n = \lambda_n n$, on déduit de la dernière inégalité, au terme de transformations élémentaires, que

$$\frac{\nu_n}{\nu_{n+1}} \geqslant \frac{n}{n+1}\left(1 + \mu\frac{\nu_{n+1}}{n+1}\right). \tag{8}$$

Si $\nu_{n+1} \geqslant 2/\mu$ pour un $n$, il résulte de (8) que $\nu_{n+1} \leqslant \nu_n$. Donc $\nu_{n+1} \leqslant \max(2/\mu, \nu_n)$, d'où il suit que $\nu_n \leqslant \max(2/\mu, \nu_1)$ pour tous les $n$. Ce qui prouve le lemme.

DÉMONSTRATION du théorème 3. Posons $\lambda_n = \Phi(x_n) - Q$. Nous admettons que $\lambda_n > 0$ pour tous les $n$. Montrons que la suite $\{\lambda_n\}$ vérifie les conditions du lemme 4 (pour un $\mu > 0$). A cet effet, nous allons évaluer d'abord la quantité $\lambda_n - \lambda_{n+1} = \Phi(x_n) - \Phi(x_{n+1})$. Le lemme 2 nous dit que pour tous les $\alpha$ réels

$$\Phi(x_n + \alpha z_{n+1}) \leqslant \Phi(x_n) + \alpha\Phi'(x_n)(z_{n+1}) + \frac{\alpha^2}{2}L\,\|z_{n+1}\|^2.$$

Comme $\|z_{n+1}\| = 1$ et $\Phi'(x_n)(z_{n+1}) = -\|\Phi'(x_n)\|$, il vient

$$\Phi(x_n + \alpha z_{n+1}) \leqslant \Phi(x_n) - \alpha\,\|\Phi' x_n\| + \frac{\alpha^2}{2}L. \tag{9}$$

Soit $\tilde{\alpha} = \frac{1}{L}\|\Phi'(x_n)\|$. (Ce point est point de minimum pour le trinôme du second degré du second membre de (9).) En utilisant (9), on obtient

$$\lambda_n - \lambda_{n+1} = \Phi(x_n) - \Phi(x_{n+1}) \geqslant \Phi(x_n) - \Phi(x_n + \tilde{\alpha} z_{n+1}) \geqslant$$
$$\geqslant \tilde{\alpha}\,\|\Phi'(x_n)\| + \frac{\tilde{\alpha}^2}{2}L = \frac{L}{2}\|\Phi'(x_n)\|^2.$$

D'autre part, en vertu du lemme 3, pour un $c > 0$ on a

$$\lambda_n = \Phi(x_n) - Q \leqslant c\,\|\Phi'(x_n)\| \quad (n = 1, 2, \ldots).$$

Donc, pour tous les $n$

$$\lambda_n - \lambda_{n+1} \geqslant \frac{L}{2c^2}\lambda_n^2.$$

La démonstration s'achève par une référence au lemme 4.

On n'a guère de chance d'améliorer la majoration de la convergence pour une vaste classe de fonctionnelles considérées dans le théorème 3. Cependant, la méthode converge bien plus vite pour les « bonnes » fonctionnelles. Une fonctionnelle deux fois différentiable, définie sur un espace hilbertien **H**, est dite fortement convexe si existent des nombres strictement positifs $M$ et $m$ tels que

$$m \| z \|^2 \leqslant (\Phi''(x)(z), z) \leqslant M \| z \|^2 \qquad (10)$$

pour tous les $x$, $z \in \mathbf{H}$. Ici *) $\Phi''(x)$ est la dérivée seconde de la fonctionnelle $\Phi$ en $x$. Il est immédiat de vérifier que la convexité forte de $\Phi$ entraîne la convexité. Les inégalités (10) expriment le fait que les opérateurs $\Phi''(x)$ sont définis positifs pour tous les $x \in \mathbf{H}$ et que leurs bornes sont uniformément limitées par des constantes strictement positives. En particulier, la fonctionnelle quadratique $F$ envisagée au § 1 est fortement convexe (pour cette fonctionnelle $F''(x) = U$, $\forall x$).

Les résultats relatifs à la méthode de la plus forte pente pour une fonctionnelle fortement convexe sont semblables sur de nombreux points aux résultats établis au § 1 pour une fonctionnelle quadratique. On peut prouver que pour tout $x_0 \in \mathbf{H}$ l'ensemble $\Omega_0 = \{z \in \mathbf{H}, \Phi(z) \leqslant \Phi(x_0)\}$ est borné et la fonctionnelle $\Phi$ atteint son minimum sur cet ensemble (donc, sur l'espace tout entier); de plus le point de minimum $x^*$ est unique. Comme pour la fonctionnelle quadratique, on a la majoration suivante de la convergence

$$\| x_n - x^* \| \leqslant C \left\| \frac{M-m}{M+m} \right\|^n,$$

où $C$ est une constante dépendant du point initial $x$ et de la fonctionnelle $\Phi$.

4.5. Le principe de la plus forte pente est très fécond dans les problèmes d'extrémum lié. Sa mise en œuvre conduit à la *méthode du gradient lié.*

Soit $\Omega$ un ensemble convexe faiblement compact dans un espace de Banach **X**. On admet que la structure de $\Omega$ est assez simple en ce sens qu'on connaît une solution du problème de minimisation sur $\Omega$ d'une fonctionnelle linéaire. Soit, par ailleurs, $\Phi$ une fonctionnelle différentiable (non linéaire), définie dans un domaine ouvert contenant $\Omega$. On demande le minimum de $\Phi$ sur $\Omega$. La méthode du gradient lié permet de construire une suite $x_0, x_1, \ldots, x_n, \ldots$ qui est souvent minimisante. Pour $x_0$ on prend un élément quelconque de $\Omega$. Si l'on connaît l'élément $x_n$, on détermine $x_{n+1}$ par la formule

$$x_{n+1} = x_n + \varepsilon_n (z_n - x_n),$$

*) $\Phi''(x)$ est un opérateur linéaire de $\mathbf{X} = \mathbf{H}$ dans $\mathbf{X}^*$, défini par la relation $\Phi'(x+h) = \Phi'(x) + \Phi''(x)(h) + o(h)$, où $\frac{o(h)}{\| h \|} \xrightarrow[h \to 0]{} 0$. Pour plus de détails voir chap. XVII.

où $z_n \in \Omega$ est choisi à partir de la condition $\Phi'(x_n)(z_n) = \min_{z\in\Omega} \Phi'(x_n)(z)$, et la valeur de la pente $\varepsilon_n$, à partir de l'égalité $\Phi(x_{n+1}) = \min_{\alpha\in[0,1]} \Phi(x_n + \alpha(z_n - x_n))$. (Il existe d'autres procédés pour déterminer la valeur de la pente, mais nous n'insistons pas sur eux.) Donc, si dans la méthode de la plus forte pente, la direction de la pente est donnée par minimisation de la dérivée $\Phi'(x)$ sur la sphère unité (la sphère des directions) et la valeur de la pente par minimisation de la fonctionnelle sur une demi-droite, dans la méthode du gradient lié la direction de la pente s'obtient par minimisation de la dérivée sur l'ensemble donné et la valeur de la pente par minimisation de la fonctionnelle sur un intervalle fermé. *Grosso modo*, la méthode du gradient lié est une modification de la méthode de la plus forte pente pour la recherche de l'extrémum sous des contraintes.

La méthode du gradient lié est justiciable de théorèmes analogues aux théorèmes 1, 2 et 3. Ceci étant, un point $y \in \Omega$ est point stationnaire de la fonctionnelle $\Phi$ sur l'ensemble $\Omega$, si $\Phi'(y)(y) = \min_{z\in\Omega} \Phi'(y)(z)$. Il est immédiat de vérifier que les points de minimum local de $\Phi$ sur $\Omega$ sont stationnaires et que les points stationnaires de $\Phi$, fonctionnelle convexe, sont les points de minimum global sur $\Omega$. Si X est un espace hilbertien, $\Phi$ une fonctionnelle convexe, $\inf_{x\in\Omega} \|\Phi'(x)\| > 0$ et $\Omega$ un ensemble fortement convexe (c'est-à-dire il existe $\gamma > 0$ tel que cet ensemble contienne avec deux quelconques de ses éléments $x$ et $z$ les éléments $\frac{x+z}{2} + \gamma u$, où $\|u\| \leqslant \|x - z\|^2$), alors la suite construite à l'aide de la méthode du gradient lié converge vers l'unique point de minimum de la fonctionnelle $\Phi$ sur $\Omega$ à la vitesse d'une progression géométrique. Un exposé plus détaillé de la méthode du gradient lié est accessible dans Démianov et Roubinov.

## § 5. Minimisation des fonctionnelles convexes dans des espaces de dimension finie

5.1. Au § 4, nous avons appliqué la méthode de la plus forte pente à la minimisation d'une fonctionnelle continûment différentiable. Plus bas, nous allons étudier la minimisation d'une fonctionnelle convexe sur un espace euclidien $\mathbf{R}^n$ de dimension finie.

Soit $\Phi$ une fonctionnelle convexe *). Il est classique que toute fonctionnelle convexe définie sur $\mathbf{R}^n$ est continue (cf. par exemple Rockafellar).

*) Voir la définition d'une fonctionnelle convexe dans le numéro 4.2.

Un vecteur $v \in \mathbf{R}^n$ est appelé *subgradient* de la fonctionnelle $\Phi$ en un point $x$ si

$$\Phi(z) \geqslant \Phi(x) + (v, \; z - x) \tag{1}$$

pour tous les $z \in \mathbf{R}^n$.

L'ensemble des subgradients n'est pas vide, est convexe, fermé et borné. On l'appelle *subdifférentielle* de la fonction $\Phi$ au point $x$ et on le note $\partial\Phi(x)$. La subdifférentielle est une application multivalente qui sera étudiée en détail au chapitre XVI, § 5. Pour l'instant on remarquera simplement que l'application $\partial\Phi(x)$ est semi-continue supérieurement, c'est-à-dire si $x_k \to x$, $v_k \to v$, $v_k \in \partial\Phi(x_k)$, alors $v \in \partial\Phi(x)$.

La fonctionnelle convexe $\Phi$ est dérivable suivant des directions, c'est-à-dire pour tous $x \in \mathbf{R}^n$ et $y \in \mathbf{R}^n$ il existe

$$\frac{1}{\|y\|} \lim_{\alpha \to +0} \frac{1}{\alpha} [\Phi(x + \alpha y) - \Phi(x)] \equiv \frac{\partial\Phi}{\partial y}(x),$$

de plus

$$\|y\| \frac{\partial\Phi}{\partial y}(x) = \max_{v \in \partial\Phi(x)} (v, \; y). \tag{2}$$

La direction de $y(x)$, $\|y(x)\| = 1$, s'appelle *direction de plus forte pente* de la fonctionnelle $\Phi$ au point $x$ si (cf. (4.1))

$$\frac{\partial\Phi}{\partial y(x)}(x) = \inf_{\|y\|=1} \frac{\partial\Phi}{\partial y}(x).$$

La relation (2) nous dit que si $\inf\limits_{\|y\|=1} \frac{\partial\Phi}{\partial y}(x) < 0$, la direction de plus forte pente existe, est unique et peut être déterminée par la formule (cf. avec l'exemple 1 du numéro 4.1)

$$y(x) = -\frac{z(x)}{\|z(x)\|}, \tag{3}$$

où

$$\|z(x)\| = \min_{z \in \partial\Phi(x)} \|z\| \equiv \rho(x).$$

LEMME 1. *Une condition nécessaire et suffisante pour qu'une fonctionnelle convexe $\Phi$ atteigne son minimum en un point $x$ de $\mathbf{R}^n$ est que*

$$\inf_{\|y\|=1} \frac{\partial\Phi}{\partial y}(x) \geqslant 0. \tag{4}$$

DÉMONSTRATION. La condition nécessaire est évidente. Prouvons la condition suffisante. Supposons par absurde que, la condition (4) étant réalisée, le point $x$ n'est pas point de minimum. Supposons qu'existe un point $z$ tel que $\Phi(z) < \Phi(x)$. Considérons la direction de $y = \frac{z-x}{\|z-x\|}$. Par définition de la fonctionnelle con-

vexe, on aurait pour de petits $\alpha$

$$\Phi(x+\alpha y)=\Phi\left(\left(1-\frac{\alpha}{\|z-x\|}\right)x+\frac{\alpha}{\|z-x\|}z\right)\leqslant$$
$$\leqslant\left(1-\frac{\alpha}{\|z-x\|}\right)\Phi(x)+\frac{\alpha}{\|z-x\|}\Phi(z),$$

donc

$$\frac{\partial\Phi}{\partial y}(x)\equiv\lim_{\alpha\to+0}\frac{1}{\alpha}[\Phi(x+\alpha y)-\Phi(x)]\leqslant\frac{1}{\|z-x\|}(\Phi(z)-\Phi(x))<0,$$

ce qui contredit (4).

La signification géométrique de la condition (4) est

$$0\in\partial\Phi(x). \tag{5}$$

La condition (5) est une généralisation de la condition nécessaire de minimum des fonctionnelles continûment dérivables: $\Phi'_x = 0$. Citons un exemple. Soit

$$\Phi(x)=\max_{z\in G}F(x, z), \tag{6}$$

où $G$ est un compact, la fonctionnelle $F(x, z)$ est continue avec $F'_x(x, z)$ par rapport à l'ensemble des variables sur $\mathbf{R}^n\times G$ et est convexe par rapport à $x$ pour tout $z\in G$ fixe. Alors $\Phi$ est convexe et de plus

$$\partial\Phi(x) = \mathrm{co}\,\{F'_x(x, z):\ z\in R(x)\},$$

où

$$R(x) = \{z\in G:\ F(x, z) = \Phi(x)\},$$

co $A$ désignant l'enveloppe convexe de l'ensemble $A$.

Remarque. La fonctionnelle (6) est généralement dérivable suivant des directions même lorsque $F$ n'est pas convexe par rapport à $x$. Bien plus, la formule (2) subsiste, quoique $\partial\Phi(x)$ ne soit déjà plus subdifférentielle.

Soit $\varepsilon>0$. Un vecteur $v$ est *ε-subgradient* d'une fonctionnelle $\Phi$ en un point $x$ si

$$\Phi(z)\geqslant\Phi(x)+(v, z-x)-\varepsilon, \tag{7}$$

pour tous les $z\in\mathbf{R}^n$.

Désignons par $\partial_\varepsilon\Phi(x)$ l'ensemble de tous les ε-subgradients. Cet ensemble est fermé, convexe et borné. Un point $x$ est un *point ε-stationnaire* d'une fonctionnelle $\Phi$ si

$$0\in\partial_\varepsilon\Phi(x). \tag{8}$$

La relation (7) nous dit alors que

$$0\leqslant\Phi(x)-\min_{z\in\mathbf{R}^n}\Phi(z)\leqslant\varepsilon. \tag{9}$$

Si $0 \notin \partial_\varepsilon \Phi(x)$, la direction de $y_\varepsilon(x) = -\frac{z_\varepsilon(x)}{\|z_\varepsilon(x)\|}$, où $\|z_\varepsilon(x)\| = \min_{z \in \partial_\varepsilon \Phi(x)} \|z\| \equiv \rho_\varepsilon(x)$, est appelée *direction de ε-plus forte pente* de la fonctionnelle $\Phi$ au point $x$. L'application $\partial_\varepsilon \Phi(x)$ est aussi semi-continue supérieurement.

Soit $v \in \partial\Phi(x)$. Alors de (1) il s'ensuit que

$$\Phi(z) \geqslant \Phi(x) + (z - x, v) = \\ = \Phi(x_0) + (v, z - x_0) + [\Phi(x) - \Phi(x_0) + (v, x_0 - x)].$$

D'où il est clair que pour $x$ assez voisin de $x_0$

$$\partial\Phi(x) \subset \partial_\varepsilon \Phi(x_0). \tag{10}$$

5.2. Il est naturel de considérer la méthode de la plus forte pente dans cette situation.

Soit $x_0 \in \mathbf{R}^n$. Supposons qu'on ait trouvé $x_k \in \mathbf{R}^n$. Si $0 \in \partial\Phi(x_k)$, le point $x_k$ est point de minimum et le processus prend fin. Si $0 \notin \partial\Phi(x_k)$, on cherche $y_{k+1} = y(x_k)$ (cf. (3)). La direction de $y_{k+1}$ est direction de plus forte pente de la fonctionnelle $\Phi$ au point $x_k$. Considérons la demi-droite

$$x_{k\alpha} = x_k + \alpha y_{k+1} \quad (\alpha \geqslant 0)$$

et trouvons

$$\min_{\alpha \geqslant 0} \Phi(x_{k\alpha}) = \Phi(x_{k\alpha_k}).$$

Posons $x_{k+1} = x_{k\alpha_k}$. Il est évident que $\Phi(x_{k+1}) < \Phi(x_k)$.

Cependant, la méthode que nous venons de décrire peut ne pas converger vers le point de minimum (effet d'« enrayage »), à cause de la discontinuité de l'application $\partial\Phi(x)$ (on trouvera un exemple dans l'ouvrage de Démianov et Malozémov).

5.3. Soit $\varepsilon > 0$. On peut se servir de la méthode suivante pour déterminer les points ε-stationnaires. Soit $x_0 \in \mathbf{R}^n$. Supposons que l'ensemble $\Omega_0 = \{x : \Phi(x) \leqslant \Phi(x_0)\}$ est borné. Supposons d'autre part qu'on ait déjà trouvé $x_k \in \mathbf{R}^n$. Si $0 \in \partial_\varepsilon \Phi(x_k)$, le point $x_k$ est ε-stationnaire, et le processus s'interrompt. Si $0 \notin \partial_\varepsilon \Phi(x_k)$, on trouve $y_{k+1} = y_\varepsilon(x_k)$ et on considère la demi-droite $x_{k\alpha} = x_k + \alpha y_{k+1}$ $(\alpha \geqslant 0)$. On cherche $\min_{\alpha \geqslant 0} \Phi(x_{k\alpha}) = \Phi(x_{k\alpha_k})$. On pose $x_{k+1} = x_{k\alpha_k}$. Il est clair que $\Phi(x_{k+1}) < \Phi(x_k)$.

En poursuivant cette procédure, on obtient une suite de points $\{x_k\}$. Si cette suite est finie, son dernier point est ε-stationnaire. Si elle est infinie, on a le

THÉORÈME 1. *Tout point limite de la suite $\{x_k\}$ est point ε-stationnaire de la fonctionnelle $\Phi$.*

11—01292

Démonstration. Soit $x_{k_s} \to \tilde{x}$. On demande de prouver que $0 \in \partial_\varepsilon \Phi(\tilde{x})$. Supposons le contraire. Soit $\rho_\varepsilon(\tilde{x}) = a > 0$. La semi-continuité supérieure de l'application $\partial_\varepsilon \Phi(x)$ entraîne l'existence d'un $\delta > 0$ tel que $\rho_\varepsilon(x) \geqslant \frac{a}{2}$ pour tous les

$$x \in S_\delta(\tilde{x}) = \{x : \|x - \tilde{x}\| \leqslant \delta\}.$$

Donc, pour les $k_s$ assez grands on aura

$$\rho_\varepsilon(x_{k_s}) \geqslant \frac{a}{2}. \tag{11}$$

De (10) il résulte qu'il existe $\delta_1 > 0$ tel que $\partial\Phi(x) \subset \partial_\varepsilon \Phi(x_{k_s})$ pour tous les $x \in S_{\delta_1}(x_{k_s})$ (et les $k_s$ assez grands). Donc, sur la demi-droite $x_{k_s \alpha}$, la fonctionnelle $\Phi$ décroît à une vitesse de valeur absolue $\geqslant a/2$ et ce à partir d'une distance $\geqslant \delta_1$. Donc

$$\Phi(x_{k_s+1}) \leqslant \Phi(x_{k_s}) - \delta_1 \frac{a}{2}.$$

D'où $\Phi(x_k) \to -\infty$, ce qui contredit le fait que l'ensemble $\Omega_0$ est borné.

5.4. Pour trouver les points de minimum de la fonctionnelle $\Phi$ on peut procéder désormais comme suit: on prend $\varepsilon_0 > 0$, $\rho_0 > 0$, $x_0 \in \mathbf{R}^n$ quelconques. On suppose déjà trouvés $\varepsilon_k > 0$, $\rho_k > 0$, $x_k \in \mathbf{R}^n$. Si $0 \in \partial\Phi(x_k)$, alors $x_k$ est point de minimum. Dans le cas contraire, en se servant de la méthode décrite au numéro 5.3 (en prenant $x_k$ pour point initial), on obtient en un nombre fini de pas un point $x_{k+1}$ tel que $\rho_{\varepsilon_k}(x_{k+1}) \leqslant \rho_k$. On pose ensuite $\varepsilon_{k+1} = \beta\varepsilon_k$, $\rho_{k+1} = \beta\rho_k$, où $\beta \in ]0, 1[$ et ne dépend pas de $k$. On construit ainsi la suite $\{x_k\}$. On établit immédiatement (cf. par exemple la démonstration du théorème III.7.2, tome 1, dans Démianov et Malozémov) le

Théorème 2. *Tout point limite de la suite $\{x_k\}$ est point de minimum de la fonctionnelle $\Phi$.*

Remarque. Nous avons décrit plus haut les schémas théoriques des algorithmes. Quand on se sert de ces algorithmes, on cherche la solution de nombreux problèmes auxiliaires (détermination de la direction de la pente, minimisation sur une demi-droite) par approximations.

5.5. Voyons maintenant la méthode du gradient généralisé (cf. Ermoliev et Chor [1]). Soient $\Phi$ une fonctionnelle convexe sur $\mathbf{R}^n$, atteignant son minimum sur $\mathbf{R}^n$, $M$ l'ensemble des points de minimum, $\rho(x) = \min_{z \in M} \|z - x\|$. Désignons par $M_\varepsilon$ l'$\varepsilon$-voisinage de $M$:

$$M_\varepsilon = \{x : \rho(x) \leqslant \varepsilon\}, \quad \varepsilon > 0.$$

Soit $\partial\Phi(x)$ la subdifférentielle de $\Phi$ en $x$:

$$\partial\Phi(x)=\{z\in\mathbf{R}^n:\ \Phi(y)-\Phi(x)\geqslant(z,\ y-x),\quad \forall y\in\mathbf{R}^n\}.$$

Choisissons $x_0\in\mathbf{R}^n$ et une suite de nombres positifs $\{\lambda_k\}$, telle que $\lambda_k\to 0$, $\sum\limits_{k=0}^{\infty}\lambda_k=\infty$.

Construisons la suite $\{x_k\}$ à l'aide des formules

$$x_{k+1}=x_k-\lambda_k y_{k+1},\quad y_{k+1}=z_{k+1}/\|z_{k+1}\|.$$

où $z_{k+1}$ est un vecteur arbitraire de $\partial\Phi(x_k)$. Si à un pas quelconque $z_k=0$, alors $x_k\in M$ et le processus s'achève.

Supposons maintenant que la suite $\{x_k\}$ est infinie.

THÉORÈME 3. *Si l'ensemble $M$ est borné, alors*

$$\rho(x_k)\to 0,\quad \Phi(x_k)\to\tilde{\Phi}\equiv\min_{x\in\mathbf{R}^n}\Phi(x).$$

DÉMONSTRATION. Soit $\Omega(x)=\{y:\ \Phi(y)\leqslant\Phi(x)\}$. L'ensemble $M$ étant borné, $\Omega(x)$ l'est aussi pour tout $x\in\mathbf{R}^n$ (cf. Rockafellar). Posons

$$T(x)=\{y:\ \Phi(y)=\Phi(x)\},\quad b(x)=\min_{y\in T(x)}\rho(y).$$

On prouve sans peine que

$$M_{b(x)}\subset\Omega(x). \tag{12}$$

Soit $x\notin M$. Alors $b(x)>0$. Considérons une direction quelconque $y=\frac{z}{\|z\|}$, $z\in\partial\Phi(x)$. Pour tous les $v\in\Omega(x)$ on a

$$(y,\ v-x)\leqslant\frac{1}{\|z\|}(\Phi(v)-\Phi(x))\leqslant 0.$$

En particulier, en vertu de (12) on a pour tous les $z\in M$

$$(y,\ z+b(x)\,y-x)\leqslant 0.$$

D'où l'on déduit sans peine que la fonctionnelle $\rho$ décroît dans la direction $-y$ et que

$$\lim_{k\to\infty}b(x_k)=0. \tag{13}$$

De (13) il résulte qu'existe une suite partielle $\{x_{k_s}\}$ telle que $b(x_{k_s})\to 0$. Donc $\rho(x_{k_s})\to 0$. En réalité, on peut prouver que toute la suite $\rho(x_k)$ converge vers 0, c'est-à-dire $\Phi(x_k)\to\tilde{\Phi}$.

REMARQUE. Si la fonctionnelle $\Phi$ est continûment dérivable, sa subdifférentielle est constituée d'un unique point (son gradient). Dans ce cas la méthode de la plus forte pente (numéro 5.4) coïncide

avec la méthode de la plus forte pente, développée au § 4, sous réserve que la valeur de la pente soit choisie d'après le deuxième procédé du numéro 1.1. La méthode du gradient généralisé (numéro 5.5) est confondue avec la méthode de la plus forte pente, sous réserve que la valeur de la pente soit déterminée par le troisième procédé du numéro 1.1.

La méthode de la plus forte pente est développée en détail dans Démianov et Malozémov, quant à la méthode du gradient généralisé, elle figure par exemple dans le travail d'Ermoliev et Chor [1].

Les méthodes décrites dans ce paragraphe ont été généralisées à la minimisation sous certaines contraintes. Cette généralisation se base sur les conditions nécessaires d'existence du minimum (cf. par exemple Doubovitski et Milioutine [1], ainsi que la monographie de Pchénytchni).

CHAPITRE XVI

# PRINCIPE DU POINT FIXE

## § 1. Principe de Caccoppoli-Banach

Ce chapitre et le suivant seront essentiellement consacrés aux opérateurs et équations non linéaires.

**1.1.** Nous commençons l'étude des équations non linéaires par un cas élémentaire généralisant le théorème de Banach de l'opérateur inverse (V.4.5, tome 1).

Soit un ensemble fermé $\Omega$ dans un espace métrique complet (pas nécessairement vectoriel) X. Soit donnée sur $\Omega$ une application $P$ de $\Omega$ dans lui-même. Un point $x^* \in \Omega$ est par définition *point fixe de l'application* $P$ si

$$x^* = P(x^*).$$

Donc, les points fixes de l'application $P$ sont solutions de l'équation

$$x = P(x). \tag{1}$$

L'application $P$ peut ne posséder aucun point fixe; tel est le cas par exemple, si $X = \Omega$ est un espace métrique vectoriel, et

$$P(x) = x + x_0 \quad (x_0 \neq 0).$$

On dira que l'application $P$ est *opérateur de contraction* si existe un nombre $\alpha < 1$ tel que

$$\rho(P(x), P(x')) \leqslant \alpha\rho(x, x') \tag{2}$$

pour tous les $x, x' \in \Omega$. Si $P$ est opérateur de contraction, il existe un point fixe et un seul. Plus exactement, on a le

THÉORÈME 1. *Si $P$ est opérateur de contraction, l'équation* (1) *admet une solution unique $x^*$ dans $\Omega$.*

*La solution $x^*$ est la limite de la suite $\{x_n\}$, où*

$$x_{n+1} = P(x_n) \quad (n = 0, 1, \ldots),$$

*et $x_0$ un élément arbitraire de $\Omega$.*

*La vitesse de convergence de la suite $\{x_n\}$ vers la solution est donnée par*

$$\rho(x_n, x^*) \leqslant \frac{\alpha^n}{1-\alpha}\rho(x_1, x_0) \quad (n = 0, 1, \ldots). \tag{3}$$

DÉMONSTRATION. Comme

$$x_{n+1} = P(x_n), \quad x_n = P(x_{n-1}),$$

on a en vertu de (2)

$$\rho(x_{n+1}, x_n) \leqslant \alpha\rho(x_n, x_{n-1}).$$

En utilisant de proche en proche la même inégalité, on obtient

$$\rho(x_{n+1}, x_n) \leqslant \alpha^n \rho(x_1, x_0).$$

Donc

$$\rho(x_{n+p}, x_n) \leqslant \rho(x_{n+p}, x_{n+p-1}) + \ldots + \rho(x_{n+1}, x_n) \leqslant$$
$$\leqslant (\alpha^{n+p-1} + \ldots + \alpha^n)\rho(x_1, x_0) \leqslant \frac{\alpha^n}{1-\alpha}\rho(x_1, x_0). \quad (4)$$

Comme $\alpha^n \to 0$, la majoration obtenue montre que la suite $\{x_n\}$ est une suite de Cauchy, donc, en vertu de la complétude de $\mathbf{X}$, elle converge vers un élément $x^* \in \mathbf{X}$. L'élément $x^* \in \Omega$, car chaque $x_n \in \Omega$ et $\Omega$ est fermé. Donc, $P(x^*)$ a un sens. En utilisant de nouveau (2), on obtient

$$\rho(x_{n+1}, P(x^*)) = \rho(P(x_n), P(x^*)) \leqslant \alpha\rho(x_n, x^*) \quad (n = 0, 1, \ldots)$$

et, comme le second membre tend vers 0,

$$x_{n+1} \to P(x^*)$$

d'où

$$x^* = P(x^*).$$

L'unicité de la solution découle également de (2). En effet, s'il existait une autre solution $\tilde{x} \in \Omega$, on aurait

$$\rho(\tilde{x}, x^*) = \rho(P(\tilde{x}), P(x^*)) \leqslant \alpha\rho(\tilde{x}, x^*),$$

ce qui n'est possible que si $\rho(\tilde{x}, x^*) = 0$, c'est-à-dire $\tilde{x} = x^*$.

On obtient la majoration (3) si dans (4) on passe à la limite pour $p \to \infty$.

REMARQUE. L'inégalité (3) définit le domaine contenant la solution de l'équation (1). En particulier, pour $n = 0$

$$\rho(x^*, x_0) \leqslant \frac{1}{1-\alpha}\rho(x_1, x_0). \quad (5)$$

1.2. D'une manière générale, la condition (2) ne peut être remplacée par une plus faible

$$\rho(P(x), P(x')) < \rho(x, x') \quad (x, x' \in \Omega, \quad x \neq x'). \quad (6)$$

En effet, soit $\mathbf{X} = \mathbf{R}^1$ l'ensemble des réels, $\Omega = \mathbf{X}$, l'application $P$ étant définie comme suit:

$$P(x) = \frac{\pi}{2} + x - \operatorname{arctg} x.$$

Il est immédiat que l'application $P$ ne possède pas de points fixes, alors que

$$\rho(P(x), P(x')) = |P(x) - P(x')| =$$
$$= |P'(\xi)|\,|x - x'| = \left|\frac{\xi^2}{1+\xi^2}\right| |x - x'| < \rho(x, x')$$

($\xi$ est compris entre $x$ et $x'$). On a toutefois le

Théorème 2. *Si l'application $P$ envoie un ensemble fermé $\Delta \subset \Omega$ dans un ensemble relativement compact $\overline{\Delta} \subset \Omega$ et si est vérifiée la condition* (6), *l'application $P$ admet un seul point fixe.*

Démonstration. Considérons l'application $P$ sur le compact $\overline{\Delta}$. La continuité évidente de $P$ entraîne celle de la fonction $\varphi(x) = \rho(x, P(x))$. La fonction $\varphi$ prend sur $\overline{\Delta}$ sa valeur minimale en $x_0 \in \overline{\Delta}$:

$$\rho(x_0, P(x_0)) = \min_{x \in \overline{\Delta}} \rho(x, P(x)).$$

Supposons que $\rho(x_0, P(x_0)) > 0$. D'après (6)

$$\rho(P(x_0), P^2(x_0)) < \rho(x_0, P(x_0)) = \min_{x \in \overline{\Delta}} \rho(x, P(x)). \tag{7}$$

Mais $P(x_0) \in \overline{\Delta}$, ce qui contredit l'inégalité (7). Donc, $\rho(x_0, P(x_0)) = 0$ et $x_0 = P(x_0)$.

Si $\tilde{x} \in \Omega$ est un autre point fixe de $P$, alors

$$\rho(\tilde{x}, x_0) = \rho(P(\tilde{x}), P(x_0)) < \rho(\tilde{x}, x_0),$$

cette contradiction prouve le théorème.

1.3. Il arrive parfois que l'application $P$ dépende d'un paramètre numérique ou autre. La solution de l'équation (1) dépendra alors aussi de ce paramètre. On peut montrer que la dépendance continue de $P$ par rapport au paramètre entraîne celle de la solution. Donnons des énoncés plus précis.

Soit $\mathbf{Y}$ un autre espace métrique. Supposons qu'à tout $y \in \mathbf{Y}$ est associée une application $P_y$ de $\Omega \subset \mathbf{X}$ dans lui-même. On dira que *$P_y$ est continue en $y$ au point $y_0 \in \mathbf{Y}_0$* si quelle que soit la suite $\{y_n\} \subset \mathbf{Y}_0$, $y_n \to y_0$,

$$P_{y_n}(x) \to P_{y_0}(x), \quad \forall x \in \Omega. \tag{8}$$

Soit l'équation (ou plus exactement la famille d'équations)

$$x = P_y(x). \tag{9}$$

On suppose que pour tout $y \in \mathbf{Y}$ elle admet une solution unique qui, de toute évidence, dépendra de $y$, et qui sera naturellement désignée par $x_y^*$. On dira que la solution de l'équation (9) *dépend continûment*

*de* $y$ *pour* $y = y_0$ si quelle que soit la suite $\{y_n\} \subset \mathbf{Y}$, $y_n \to y_0$, on aura $x^*_{y_n} \to x^*_{y_0}$.

Si l'application $P_y$ est opérateur de contraction pour tout $y \in \mathbf{Y}_0$, sa continuité entraîne celle de la solution de (9). Plus exactement on a le

THÉORÈME 3. *Si pour tout* $y \in \mathbf{Y}_0$ *l'application* $P_y$ *vérifie la condition* (2) *avec* $\alpha$ *ne dépendant pas de* $y$ *et qu'elle soit continue en* $y$ *au point* $y_0 \in \mathbf{Y}_0$, *alors pour* $y = y_0$ *la solution de l'équation* (9) *dépend continûment de* $y$.

DEMONSTRATION. Soit $y$ un élément arbitraire de $\mathbf{Y}_0$. Construisons la solution $x^*_y$ de l'équation (9) comme la limite de $\{x_n\}$:

$$x_{n+1} = P_y(x_n) \quad (n = 0, 1, \ldots; \; x_0 = x^*_{y_0}).$$

Comme $x^*_{y_0} = P_{y_0}(x^*_{y_0})$, on obtient en vertu de (5)

$$\rho(x^*_y, x^*_{y_0}) \leqslant \frac{1}{1-\alpha}\rho(x_1, x_0) = \frac{1}{1-\alpha}\rho(P_y(x^*_{y_0}), P_{y_0}(x^*_{y_0})).$$

La continuité souhaitée de $x^*_y$ pour $y = y_0$ s'établit maintenant sans peine à l'aide de (8).

REMARQUE. L'ensemble des applications $P_y$ peut être traité comme une seule application $P$ associant à tout couple $(x, y)$ d'éléments $x \in \Omega$, $y \in \mathbf{Y}_0$ l'élément $P(x, y) = P_y(x)$. De façon analogue, la solution $x^*_y$ de l'équation (9) peut être considérée comme une application $F$ associant à un élément $y \in \mathbf{Y}_0$ l'élément $F(y) = x^*_y$.

Le théorème 3 s'énonce encore:

Si pour chaque $y \in \mathbf{Y}_0$

$$\rho(P(x, y), P(x', y)) \leqslant \alpha\rho(x, x') \quad (x, x' \in \Omega)$$

($\alpha < 1$ ne dépend pas de $y$) et pour tout $x \in \Omega$ l'application $P$ est continue en $y$ au point $y_0 \in \mathbf{Y}$, alors l'application $F$ est aussi continue au point $y_0$.

Tel qu'il a été formulé, le principe du point fixe a été établi par Banach [1] et Caccoppoli. Pour les généralisations aux applications multivalentes voir Ioffe et Tikhomirov.

## § 2. Propositions auxiliaires

Dans le paragraphe suivant nous établirons le deuxième principe du point fixe. Sa démonstration pose des difficultés notoires, car elle repose sur des résultats délicats concernant la structure topologique d'un espace de dimension finie. Ces propositions auxiliaires de topologie feront l'objet du présent paragraphe.

2.1. Introduisons d'abord quelques nouvelles notions indispensables pour la suite. Soient un $B$-espace $\mathbf{X}$ et la suite de ses éléments

$$x_0, x_1, x_2, \ldots, x_n \tag{1}$$

Supposons que les différences

$$x_1 - x_0,\ x_2 - x_0,\ \ldots,\ x_n - x_0 \tag{2}$$

sont linéairement indépendantes, et formons l'enveloppe convexe $S(x_0, x_1, \ldots, x_n)$ des éléments (1). L'ensemble $S(x_0, x_1, \ldots, x_n)$ est un *simplexe*, les points $x_0, x_1, \ldots, x_n$, ses *sommets*. Le nombre $n$ est la *dimension* du simplexe.

Remarque. Dans cette définition du simplexe, les sommets ne jouent pas tous le même rôle. Cependant, on peut prouver (ce qu'on laisse au soin du lecteur) que si les différences (2) sont linéairement indépendantes, les $n$ différences restantes

$$x_0 - x_k,\ \ldots,\ x_{k-1} - x_k,\quad x_{k+1} - x_k,\ \ldots,\ x_n - x_k \qquad (k = 1, 2, \ldots, n)$$

le sont nécessairement. Ceci rétablit l'identité des rôles de tous les sommets.

Soit donné un simplexe $S(x_0, x_1, \ldots, x_n)$; considérons $k$ sommets $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ et construisons un simplexe sur les $n-k$ sommets restants *). Le simplexe obtenu s'appelle *face* (de dimension $n-k$) du simplexe $S(x_0, x_1, \ldots, x_n)$, opposée aux sommets $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$.

Soit encore un simplexe $S = S(x_0, x_1, \ldots, x_n)$. Tout élément $x \in S$ peut être mis sous la forme

$$x = \alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_n x_n, \tag{3}$$

où les coefficients $\alpha_k$ sont reliés par la relation

$$\alpha_0 + \alpha_1 + \ldots + \alpha_n = 1, \quad \alpha_k \geqslant 0 \quad (k = 0, 1, \ldots, n).$$

Prouvons que la représentation (3) est unique. En effet, comme $\alpha_0 = 1 - \sum_{k=1}^{n} \alpha_k$, il résulte de (3) que

$$x = x_0 + \sum_{k=1}^{n} \alpha_k (x_k - x_0). \tag{4}$$

S'il existe une autre représentation

$$x = \alpha'_0 x_0 + \alpha'_1 x_1 + \ldots + \alpha'_n x_n,$$

on trouve comme plus haut que

$$x = x_0 + \sum_{k=1}^{n} \alpha'_k (x_k - x_0),$$

*) Ceci est possible d'après la remarque faite plus haut.

ce qui avec (4) donne

$$\sum_{k=1}^{n} \alpha_k (x_k - x_0) = \sum_{k=1}^{n} \alpha'_k (x_k - x_0),$$

égalité qui, vu l'indépendance linéaire des différences (2), n'est possible que si

$$\alpha_1 = \alpha'_1, \ \alpha_2 = \alpha'_2, \ \ldots, \ \alpha_n = \alpha'_n, \ \alpha'_0 = \alpha_0.$$

Les nombres $\alpha_0, \alpha_1, \ldots, \alpha_n$ sont appelés *coordonnées simpliciales* du point $x \in S$ *).

On remarquera que $\alpha_{i_1} = \alpha_{i_2} = \ldots = \alpha_{i_k} = 0$ pour les points de la face opposée aux sommets $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$.

On appelle *point intérieur* d'un simplexe tout point de coordonnées positives.

On appelle *centre* d'un simplexe un point dont toutes les coordonnées sont égales.

Introduisons la notion de subdivision d'un simplexe $S = S(x_0, x_1, \ldots, x_n)$. On appelle *subdivision* d'un simplexe sa décomposition en une somme d'un nombre fini de simplexes de même dimension que lui, définis par récurrence de la manière suivante.

Pour subdivision du simplexe à une dimension $S(x_0, x_1)$ on convient de prendre l'ensemble des deux simplexes $S\left(x_0, \frac{1}{2}(x_0 + x_1)\right)$ et $S\left(\frac{1}{2}(x_0 + x_1), x_1\right)$. Supposons qu'on connaît la subdivision de tous les simplexes de dimension inférieure à $n$ et considérons le simplexe de dimension $n$ $S = S(x_0, x_1, \ldots, x_n)$. Soit une face de dimension $n - 1$ de $S$. Cette face étant un simplexe de dimension $n - 1$, on connaît déjà sa subdivision. Supposons qu'elle est constituée des simplexes $(n - 1)$-dimensionnels $S_1, S_2, \ldots, S_m$. En ajoutant aux sommets d'un simplexe $S_k$ $(k = 1, 2, \ldots, m)$ le centre du simplexe $S$, on obtient un simplexe de dimension $n$ **). On obtient une subdivision du simplexe en effectuant les mêmes constructions pour chaque face de dimension $n - 1$.

Il est aisé de voir que les simplexes d'une subdivision ne possèdent pas de points intérieurs communs. Ils se coupent suivant la face commune.

Considérons un simplexe $S$ et une de ses subdivisions. En subdivisant chaque simplexe partiel, on obtient une double subdivision de $S$. En poursuivant cette opération, on obtient des subdivisions de multiplicité plus élevée.

*) Les nombres $\alpha_0, \alpha_1, \ldots, \alpha_n$ sont parfois appelés *coordonnées barycentriques* du point $x \in S$, car si l'on place une masse $\alpha_k$ en chaque sommet $x_k$, le centre d'inertie du système de points matériels obtenu sera en $x$.

**) On laisse au lecteur le soin de vérifier que cette construction est possible.

Il est clair que le diamètre des simplexes tend vers zéro avec l'accroissement de la multiplicité de la subdivision.

**2.2.** Indiquons maintenant les lemmes qui nous serviront pour démontrer le deuxième principe du point fixe.

Soient le simplexe $S = S(x_0, x_1, \ldots, x_n)$ et une de ses subdivisions (de multiplicité quelconque), composée des simplexes

$$S_1, S_2, \ldots, S_m. \tag{5}$$

Supposons qu'à chaque sommet $z$ de chacun des simplexes (5) est associé l'un des nombres $0, 1, \ldots, n$, c'est-à-dire sur l'ensemble des sommets des simplexes (5) est définie une fonction $\nu$ prenant les valeurs $0, 1, \ldots, n$. Donc, à chaque simplexe (5) est associé l'ensemble $\nu(S_j)$ des $n + 1$ nombres correspondant aux sommets de ce simplexe. On dira qu'un simplexe $S_j$ est *normal* si $\nu(S_j) = \{0, 1, \ldots, n\}$ indépendamment de l'ordre dans lequel sont disposés les nombres de $\nu(S_j)$.

Si la fonction $\nu$ est arbitraire, il peut n'exister aucun simplexe normal. Cependant on a le

LEMME 1. *Supposons que la fonction $\nu$ vérifie la condition suivante: si $z$ est un sommet d'un simplexe de* (5), *appartenant à une face $k$-dimensionnelle $S(x_{i_0}, x_{i_1}, \ldots, x_{i_k})$ de $S$, alors $\nu(z)$ ne peut prendre que l'une des valeurs $i_0, i_1, \ldots, i_k$.*

*Alors parmi les simplexes* (5) *il en existe au moins un normal.*

DÉMONSTRATION. Nous allons prouver que le nombre de simplexes normaux est impair. Si le simplexe $S$ est de dimension zéro, c'est-à-dire si $S$ est un point, le lemme est trivial. Raisonnons par récurrence. Supposons que le lemme est vrai pour les simplexes de dimension $n - 1$ et prouvons-le pour le simplexe de dimension $n$. Soit $S_k^{(i)}$ une face de dimension $n - 1$ du simplexe $S_k$. On dira qu'elle est *distinguée* si $\nu(S_k^{(i)}) = \{0, 1, 2, \ldots, n\}$. Si le simplexe $S_k$ est normal, on s'assure immédiatement que $\rho_k$, le nombre de ses faces distinguées, est égal à un; dans le cas contraire $\rho_k = 2$. Donc, la somme $\rho = \sum_{k=1}^{m} \rho_k$ est de même parité que le nombre de tous les simplexes normaux. Répartissons les faces distinguées des simplexes (5) en deux groupes. Le premier groupe sera composé des faces qui ne sont contenues dans aucune face de dimension $n - 1$ de $S$. Soit $\rho'$ le nombre de ces faces. Chaque face distinguée de ce groupe étant face commune d'exactement deux simplexes de (5), elle doit être comptée deux fois dans la somme $\rho$. Le deuxième groupe sera constitué des faces distinguées qui sont entièrement contenues dans une face de dimension $n - 1$ de $S$. Ces faces étant chacune face d'un seul simplexe de (5), elles seront en nombre $\rho'' = \rho - 2\rho'$. En utilisant l'hypothèse du lemme, on s'assure sans peine que toutes les faces

distinguées du second groupe sont contenues dans la face $S' = S(x_0, x_1, \ldots, x_{n-1})$ du simplexe $S$. Or, les faces de dimension $n - 1$ des simplexes (5), contenues dans $S'$, forment une subdivision de $S'$, et de plus la fonction $\nu$ envisagée uniquement pour cette subdivision vérifie l'hypothèse du lemme. Donc, d'après l'hypothèse de la récurrence, le nombre des simplexes normaux de cette subdivision est impair. En remarquant que les simplexes normaux sont précisément dans ce cas les faces distinguées du second groupe, on déduit que $\rho''$, et par suite $\rho = 2\rho' + \rho''$, est impair.

Le lemme est prouvé.

LEMME 2. *$S = S(x_0, x_1, \ldots, x_n)$ étant un simplexe dans un B-espace de dimension $n$, si existent des ensembles fermés $F_0, F_1, \ldots, F_n$ tels que*

$$S(x_{i_0}, x_{i_1}, \ldots, x_{i_k}) \subset \bigcup_{m=0}^{k} F_{i_m},$$

*quelle que soit la collection $i_0, i_1, \ldots, i_k$ $(k=0, 1, \ldots, n)$, alors l'intersection $\bigcap_{k=0}^{n} F_k$ n'est pas vide.*

DEMONSTRATION. Soit une subdivision de $S$ de multiplicité quelconque. Soient $z$ un sommet d'un simplexe de la subdivision, $S(x_{i_0}, x_{i_1}, \ldots, x_{i_k})$ la face contenant $z$. Par hypothèse, $z \in \bigcup_{m=0}^{k} F_{i_m}$. Supposons que $z \in F_{i_s}$ et posons $\nu(z) = i_s$. Il est clair que la fonction $\nu$ vérifie la condition du lemme 1 et, par suite, en vertu de ce dernier il existe un simplexe partiel normal dont l'intersection avec chacun des ensembles $F_j$ n'est pas vide (contient un sommet).

Considérons maintenant une suite de subdivisions de multiplicité $p$ $(p = 1, 2, \ldots)$. Pour la $p$-uple subdivision il existe des points $z_0^{(p)}, z_1^{(p)}, \ldots, z_n^{(p)}$, sommets d'un simplexe partiel normal, tels que

$$z_k^{(p)} \in F_k \quad (k=0, 1, \ldots, n; \quad p=1, 2, \ldots). \tag{6}$$

L'ensemble $S$ étant compact, on peut exhiber une suite strictement croissante $\{p_j\}$ d'entiers naturels, telle que la suite $\{z_0^{(p_j)}\}$ soit convergente. La limite de cette suite, le point $z^*$, appartient à $S$, car $S$ est fermé. Comme les diamètres des simplexes partiels tendent vers 0 pour $p \to \infty$, on a

$$z_k^{(p_j)} \to z^* \quad (k=0, 1, 2, \ldots, n)$$

D'où, en vertu de (6), et puisque $F_k$ sont fermés,

$$z^* \in F_k \quad (k = 0, 1, 2, \ldots, n).$$

c.q.f.d.

**2.3.** Le lemme suivant nous fournit déjà un principe du point fixe (le principe de Brouwer) mais à des conditions très restrictives.

LEMME 3. *Toute application continue $P$ d'un simplexe $n$-dimensionnel $S$ d'un $B$-espace $X$ dans lui-même possède un point fixe.*

DÉMONSTRATION. Soient $\alpha_0, \alpha_1, \ldots, \alpha_n$ les coordonnées simpliciales d'un point $x \in S$. Désignons par

$$\beta_j = f_j(\alpha_0, \alpha_1, \ldots, \alpha_n) \quad (j = 0, 1, \ldots, n)$$

les coordonnées simpliciales du point $P(x)$. Cette notation a un sens, puisque $P(x) \in S$. On a de toute évidence

$$\beta_0 + \beta_1 + \ldots + \beta_n = 1, \quad \beta_j \geqslant 0 \quad (j = 0, 1, \ldots, n).$$

Il est aisé de vérifier en utilisant la continuité de l'application $P$ que les fonctions $f_j$ sont continues. Désignons par $F_j$ ($j = 0, 1, 2, \ldots, n$) l'ensemble des points $x \in S$, tels que

$$\alpha_j \geqslant \beta_j.$$

Les fonctions $f_j$ étant continues, les ensembles $F_j$ sont fermés. Assurons-nous qu'elles vérifient les conditions du lemme 2.

En effet, soit $x \in S(x_{i_0}, x_{i_1}, \ldots, x_{i_k})$. Supposons que $x \notin \bigcup_{m=0}^{k} F_{i_m}$. Cela signifie qu'entre les coordonnées simpliciales $\alpha_0, \alpha_1, \ldots, \alpha_n$ et $\beta_0, \beta_1, \ldots, \beta_n$ des points $x$ et $P(x)$ respectivement, on a la relation

$$\alpha_{i_m} < \beta_{i_m} \quad (m = 0, 1, \ldots, k).$$

Comme pour $i \neq i_m$ ($m = 0, 1, \ldots, k$) on doit avoir $\alpha_i = 0$, donc $\alpha_i \leqslant \beta_i$, on obtient

$$\sum_{i=0}^{n} \alpha_i < \sum_{i=0}^{n} \beta_i,$$

ce qui est impossible, puisque les deux sommes doivent être égales à l'unité.

D'après le lemme 2, il existe un point $x^* \in \bigcap_{k=0}^{n} F_k$. Si $\alpha_0^*, \alpha_1^*, \ldots, \alpha_n^*$ et $\beta_0^*, \beta_1^*, \ldots, \beta_n^*$ sont les coordonnées simpliciales des points $x^*$ et $P(x^*)$ respectivement, alors

$$\alpha_j^* \geqslant \beta_j^* \quad (j = 0, 1, \ldots, n),$$

d'où

$$\sum_{i=0}^{n} \alpha_i^* \geqslant \sum_{i=0}^{n} \beta_i^*, \tag{7}$$

le signe d'égalité ayant lieu si et seulement si

$$\alpha_i^* = \beta_i^* \quad (i = 0, 1, \ldots, n). \tag{8}$$

Or, les deux sommes de (7) sont égales à un, donc (8) est vraie, c'est-à-dire $x^* = P(x^*)$.

Le lemme est prouvé.

2.4. Généralisons maintenant le lemme 3 à des ensembles convexes bornés fermés d'un espace de dimension finie. Pour cela il nous faut d'abord prouver un lemme.

On rappelle (III.2.3, tome 1) qu'un corps convexe dans un $B$-espace X est un ensemble fermé convexe $\Omega$ dont l'intérieur n'est pas vide.

LEMME 4. *Soient $\Omega$ un corps convexe borné en norme dans un B-espace* X, *$B_X$ la boule unité fermée de* X. *Les ensembles $\Omega$ et $B_X$ sont homéomorphes.*

DEMONSTRATION. Sans nuire à la généralité on peut admettre que $\mathbf{0}$ est un point intérieur de $\Omega$ et que $B_X \subset \Omega$. L'ensemble $\Omega$ est alors voisinage de $\mathbf{0}$ et on peut envisager sa fonctionnelle de Minkowski $p$ (II.3.2, tome 1).

Les propriétés générales de la fonctionnelle de Minkowski prouvées dans le lemme III.2.1 (tome 1) nous disent que $p$ est continue et $\Omega = \{x \in X : p(x) \leqslant 1\}$. Si $x \neq \mathbf{0}$ est un élément quelconque de X, alors $\frac{x}{\|x\|} \in B_X \subset \Omega$ et par suite $p\left(\frac{x}{\|x\|}\right) \leqslant 1$, ce qui conduit à l'inégalité

$$p(x) \leqslant \|x\|, \tag{9}$$

qui est visiblement valable aussi pour $x = \mathbf{0}$.

L'ensemble $\Omega$ étant borné, il existe une boule ouverte $K_r$ centrée en $\mathbf{0}$ de rayon $r > 0$, contenant l'ensemble $\Omega$. Comme $\frac{rx}{\|x\|} \notin K_r$, $\forall x \neq \mathbf{0}$, on aura également $\frac{rx}{\|x\|} \notin \Omega$, de sorte que $p\left(\frac{rx}{\|x\|}\right) > 1$. Donc

$$p(x) \geqslant \frac{1}{r}\|x\| \quad (x \in X). \tag{10}$$

Définissons l'application $T$ par

$$T(x) = x\frac{p(x)}{\|x\|} \quad (x \in \Omega,\ x \neq \mathbf{0}), \quad T(\mathbf{0}) = \mathbf{0}. \tag{11}$$

Comme l'ensemble $\Omega$ est confondu avec l'ensemble des $x \in X$ pour lesquels $p(x) \leqslant 1$, il est clair que $T(\Omega) = B_X$. De plus, il est immédiat que l'application $T$ est bijective, l'application inverse

$T^{-1}$ étant définie par

$$T^{-1}(x)=x\frac{\|x\|}{p(x)}\quad (x\neq 0),\quad T^{-1}(0)=0. \tag{12}$$

La continuité de l'application $T$ au point $x_0\neq 0$ découle de celle de la fonctionnelle $p$. Si $x_n\to 0$, alors $T(x_n)\to 0=T(0)$, puisque $\left\|\frac{x_n}{\|x_n\|}\right\|=1$ et $p(x_n)\to 0$.

On vérifie de façon analogue la continuité de l'application $T^{-1}$. Si $x_n\to x_0\neq 0$, la relation $T^{-1}(x_n)\to T^{-1}(x_0)$ est une conséquence de la continuité de la fonctionnelle $p$, puisque $p(x_0)\geqslant\frac{1}{r}\|x\|>0$ (cf. (10)). Pour $x_0=0$ on a toujours d'après (10)

$$\|T^{-1}(x_n)\|=\frac{\|x_n\|}{p(x_n)}\|x_n\|\leqslant\frac{r\|x_n\|}{\|x_n\|}\|x_n\|=r\|x_n\|\to 0,$$

c'est-à-dire $T^{-1}(x_n)\to 0=T^{-1}(0)$.

Ceci achève la démonstration du lemme.

COROLLAIRE. *Un corps convexe borné $\Omega$ d'un B-espace de dimension $n$ est homéomorphe à un simplexe de dimension $n$.*

En effet, un simplexe de dimension $n$ est un corps convexe dans un espace de dimension $n$, donc comme $\Omega$ il est homéomorphe à la boule $B$.

Maintenant nous sommes en mesure de prouver le lemme fondamental.

LEMME 5. *Soit $\Omega$ un corps convexe borné d'un B-espace* X *de dimension $n$. Une application continue $P$ de $\Omega$ dans lui-même possède un point fixe.*

DEMONSTRATION. Considérons dans l'espace X des éléments linéairement indépendants $x_1, x_2, \ldots, x_n$. Posons $x_0=0$ et considérons le simplexe $S$ de sommets $x_0, x_1, \ldots, x_n$. D'après le corollaire du lemme 4, il existe une application biunivoque et bicontinue $T$ du simplexe $S$ sur le corps $\Omega$. Soit l'application

$$P_0=T^{-1}PT.$$

L'application $P_0$ est continue et applique le simplexe $S$ dans lui-même. Donc, d'après le lemme 3, il existe un point fixe $x_0$:

$$x_0=P_0(x_0).$$

Et $x^*=T(x_0)$ sera point fixe de l'application $P$, puisque

$$P(x^*)=PT(x_0)=TT^{-1}PT(x_0)=TP_0(x_0)=T(x_0)=x^*.$$

Le théorème du point fixe d'une application continue d'un simplexe dans lui-même a été prouvé par Brouwer [1] dans le cadre de l'établissement de l'invariance de la dimension. La démonstration produite est due à Kuratowski.

## § 3. Principe de Schauder

**3.1.** A la lumière des résultats du paragraphe précédent, on peut établir le deuxième principe du point fixe (ou principe de Schauder).

THÉORÈME 1. *Toute application continue P d'un ensemble compact convexe $\Omega$ d'un B-espace X dans lui-même possède un point fixe.*

DÉMONSTRATION. Prenons un nombre arbitraire $\varepsilon > 0$. Comme $\Omega$ est un ensemble compact, il contient un $\varepsilon$-réseau fini. Supposons qu'il soit constitué des éléments

$$x_1, x_2, \ldots, x_m. \tag{1}$$

Formons l'ensemble $\Omega_0$, enveloppe convexe des éléments (1). Il est clair que $\Omega_0 \subset \Omega$. De plus, l'ensemble $\Omega_0$ est de dimension finie $n \leqslant m - 1$ *).

En raisonnant par récurrence sur la dimension, on montre sans peine que l'on peut représenter l'ensemble $\Omega_0$ sous forme d'une somme de simplexes de dimension $n$, tels que

a) tous les points (1) soient sommets de ces simplexes;

b) deux simplexes ou bien ne possèdent pas de points communs, ou bien se coupent suivant une face commune (de dimension $k$; $k = 0, 1, \ldots, n - 1$).

Procédons à une subdivision de chaque simplexe d'une multiplicité assez grande pour que les diamètres de tous les simplexes partiels soient $< \varepsilon$. Désignons ces simplexes par

$$S_1, S_2, \ldots, S_p. \tag{2}$$

Les sommets des simplexes (2) constituent visiblement un $\varepsilon$-réseau. Signalons que les simplexes (2) vérifient les conditions a) et b).

Considérons maintenant l'application $P$. L'ensemble $\Omega$ étant compact, l'application $P$ est uniformément continue, car elle est continue, c'est-à-dire pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que $\| x - x' \| < \delta$ entraîne

$$\| P(x) - P(x') \| < \varepsilon \quad (x, x' \in \Omega). \tag{3}$$

Quitte à subdiviser les simplexes (2), on peut rendre leurs diamètres plus petits que $\delta$. On suppose que cette subdivision a déjà été effectuée, c'est-à-dire les diamètres des simplexes (2) sont non seulement $< \varepsilon$, mais $< \delta$.

---

*) On dit que l'ensemble $E \subset X$ est de *dimension n* si existent des éléments $\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_n$ tels que 1) les différences $\bar{x}_k - \bar{x}_0$ ($k = 1, 2, \ldots, n$) soient linéairement indépendantes et 2) tout $x \in E$ se représente sous la forme $x = \bar{x}_0 + \sum_{k=1}^{n} \alpha_k (\bar{x}_k - \bar{x}_0)$ et de plus pour tout $k = 1, 2, \ldots, n$ il existe $x \in E$ tel que le coefficient correspondant $\alpha_k \neq 0$.

Construisons maintenant l'application $P_\varepsilon$, approximation simpliciale de l'application $P$, de $\Omega_0$ dans $\Omega_0$. Définissons tout d'abord $P_\varepsilon$ aux sommets des simplexes (2). Soit $z$ un sommet d'un simplexe (2). Comme $P(z) \in \Omega$ et que les sommets des simplexes (2) forment un ε-réseau, il existe un sommet $\bar{z}$ tel que $\| \bar{z} - P(z) \| < \varepsilon$. Posons $\bar{z} = P_\varepsilon(z)$.

Supposons maintenant que $x \in \Omega_0$ n'est sommet d'aucun simplexe (2). Soit $x \in S_k$. Désignons les sommets du simplexe $S_k$ par $x_0^{(k)}, x_1^{(k)}, \ldots, x_n^{(k)}$ et mettons l'élément $x$ sous la forme

$$x = \sum_{i=0}^{n} \alpha_i^{(k)} x_i^{(k)} \quad \Big( \sum_{i=0}^{n} \alpha_i^{(k)} = 1 ; \quad \alpha_i^{(k)} \geqslant 0 ; \quad i = 0, 1, \ldots, n \Big).$$

Posons

$$P_\varepsilon(x) = \sum_{i=0}^{n} \alpha_i^{(k)} P_\varepsilon(x_i^{(k)}). \tag{4}$$

Cette définition se passe d'explications si le simplexe contenant le point $x$ est unique. Si de plus $x \in S_r$ $(r \neq k)$, il faut encore prouver que $P_\varepsilon(x)$ ne dépend pas du simplexe choisi. D'après la condition b), l'intersection des simplexes $S_k$ et $S_r$ est leur face commune. Supposons que cette face est engendrée par les sommets $x_{i_0}^{(k)}, x_{i_1}^{(k)}, \ldots$ $\ldots, x_{i_s}^{(k)}$ du simplexe $S_k$; on peut admettre que

$$x_{i_j}^{(k)} = x_{i_j}^{(r)} \quad (j = 1, 2, \ldots, s), \tag{5}$$

$x_0^{(r)}, x_1^{(r)}, \ldots, x_n^{(r)}$ étant les sommets du simplexe $S_r$. Mettons $x$ sous la forme

$$x = \sum_{i=1}^{n} \alpha_i^{(r)} x_i^{(r)} \Big( \sum_{i=1}^{n} \alpha_i^{(r)} = 1 ; \quad \alpha_i^{(r)} \geqslant 0, \quad i = 0, 1, \ldots, n \Big).$$

Compte tenu de la remarque faite dans 2.1, on a

$$\alpha_i^{(k)} = \alpha_i^{(r)} = 0 \quad (i \neq i_j ; \quad j = 1, 2, \ldots, s). \tag{6}$$

Et en vertu de (5)

$$\sum_{j=0}^{s} \alpha_{i_j}^{(k)} x_{i_j}^{(k)} = \sum_{j=0}^{s} \alpha_{i_j}^{(r)} x_{i_j}^{(k)},$$

d'où

$$\alpha_{i_j}^{(k)} = \alpha_{i_j}^{(r)} \quad (j = 0, 1, \ldots, s). \tag{7}$$

Si donc nous définissons $P_\varepsilon(x)$ à partir du simplexe $S_r$, on obtient en vertu de (6), (5) et (7)

$$P_\varepsilon(x) = \sum_{i=0}^{n} \alpha_i^{(r)} P_\varepsilon(x_i^{(r)}) = \sum_{j=0} \alpha_{i_j}^{(r)} P_\varepsilon(x_{i_j}^{(r)}) =$$

$$= \sum_{j=0} \alpha_i^{(k)} P_\varepsilon(x_{i_j}^{(k)}) = \sum_{i=0}^{n} \alpha_i^{(k)} P_\varepsilon(x_i^{(k)}),$$

c.q.f.d.

Par des raisonnements peu compliqués, on s'assure que l'application $P_\varepsilon$ est une application continue de $\Omega_0$ dans $\Omega_0$.

Donc, l'application $P_\varepsilon$ satisfait aux conditions du lemme 5 du précédent paragraphe, par suite elle admet un point fixe $x_\varepsilon \in \Omega_0$:

$$x_\varepsilon = P_\varepsilon(x_\varepsilon).$$

Soient $z_0, z_1, \ldots, z_n$ les sommets du simplexe (2) contenant le point $x_\varepsilon$. Puisque $\| z_i - z_j \| < \delta$, on a en vertu de (3)

$$\| P(z_i) - P(z_j) \| < \varepsilon \quad (i, j = 0, 1, \ldots, n).$$

Comme par définition

$$\| P_\varepsilon(z_i) - P(z_i) \| < \varepsilon \quad (i = 0, 1, \ldots, n), \tag{8}$$

on déduit de la relation précédente que

$$\| P_\varepsilon(z_i) - P_\varepsilon(z_j) \| < 3\varepsilon \quad (i, j = 0, 1, \ldots, n). \tag{9}$$

Mettons $x_\varepsilon$ sous la forme

$$x_\varepsilon = \sum_{i=0}^{n} \alpha_i^{(\varepsilon)} z_i \quad \left(\sum_{i=0}^{n} \alpha_i^{(\varepsilon)} = 1; \quad \alpha_i^{(\varepsilon)} \geqslant 0; \quad i = 0, 1, \ldots, n\right),$$

par définition de $P_\varepsilon$ on a

$$x_\varepsilon = P_\varepsilon(x_\varepsilon) = \sum_{i=0}^{n} \alpha_i^{(\varepsilon)} P_\varepsilon(z_i).$$

D'où il suit compte tenu de (9)

$$\| x_\varepsilon - P_\varepsilon(z_j) \| = \left\| \sum_{i=0}^{n} \alpha_i^{(\varepsilon)} [P_\varepsilon(z_i) - P_\varepsilon(z_j)] \right\| \leqslant 3\varepsilon \sum_{i=0} \alpha_i^{(\varepsilon)} = 3\varepsilon \tag{10}$$
$$(j = 0, 1, \ldots, n).$$

Enfin la relation

$$\| x_\varepsilon - z_i \| < \delta \quad (i = 0, 1, \ldots, n)$$

entraîne d'après (3)

$$\| P(x_\varepsilon) - P(z_i) \| < \varepsilon \quad (i = 0, 1, \ldots, n). \tag{11}$$

En se servant de (10), (8) et (11), on trouve en définitive

$$\| x_\varepsilon - P(x_\varepsilon) \| \leqslant \| x_\varepsilon - P_\varepsilon(z_i) \| + \\ + \| P_\varepsilon(z_i) - P(z_i) \| + \| P(z_i) - P(x_\varepsilon) \| \leqslant 5\varepsilon.$$

Donc, pour tout $\varepsilon > 0$ existe un point $x_\varepsilon \in \Omega$ tel que

$$\| x_\varepsilon - P(x_\varepsilon) \| < 5\varepsilon.$$

Prenons une suite $\varepsilon_k \to 0$ et construisons pour chaque $k = 1, 2, \ldots$ le point $x_k = x_{\varepsilon_k}$. L'ensemble $\Omega$ étant compact et fermé, on peut, sans restreindre la généralité, considérer que

$$x_k \xrightarrow[k\to\infty]{} x^* \in \Omega.$$

Mais alors

$$\| x^* - P(x^*) \| \leqslant \| x^* - x_k \| + \| x_k - P(x_k) \| + \\ + \| P(x_k) - P(x^*) \|,$$

et tous les termes du second membre tendent vers zéro (le dernier en raison de la continuité de $P$). Donc,

$$x^* = P(x^*),$$

ce qui prouve l'existence du point fixe.

Remarque. On laisse au lecteur le soin de s'assurer que les deux conditions du théorème, la convexité et la compacité de $\Omega$, sont essentielles. De même, il est aisé de produire un exemple montrant que l'on ne peut affaiblir la condition de compacité de $\Omega$ et exiger en échange qu'il soit borné.

3.2. On peut énoncer le théorème 1 sous une forme plus générale en remarquant que l'enveloppe convexe fermée d'un ensemble relativement compact dans un $B$-espace est compacte (corollaire du théorème III.2.3, tome 1).

Théorème 2. *Une application continue $P$ d'un ensemble convexe fermé $\Omega$ d'un $B$-espace $X$ dans un ensemble compact $\Delta \subset \Omega$ admet un point fixe.*

Démonstration. Soit $\Omega_0 = \overline{\text{co}}\,(\Delta)$ l'enveloppe convexe fermée de $\Delta$. $\Omega_0$ est un ensemble compact convexe, et $\Omega_0 \subset \Omega$.

L'application $P$ envoie $\Omega_0$ dans $\Delta$, donc dans $\Omega_0$, puisque $\Delta \subset \Omega_0$. Par suite, si l'on envisage $P$ uniquement sur $\Omega_0$, on se retrouve dans les conditions du théorème 1, ce qui prouve le théorème 2.

En modifiant en conséquence les démonstrations des théorèmes 1 et 2, on peut établir des théorèmes analogues d'existence du point fixe dans le cas où X est un espace localement convexe (cf. Dunford et Schwartz [I]).

Le principe du point fixe (théorèmes 1 et 2) est établi dans l'article de Schauder [1]. Des théorèmes plus profonds d'existence du point fixe, en particu-

lier, pour les espaces localement convexes ont été prouvés par Leray (cf. Leray et Schauder [1]). De grands progrès ont été réalisés ces derniers temps dans l'étude de ces problèmes sous l'impulsion de M. Krasnosselski (cf. Krasnosselski [I], Krasnosselski [1], [2], [3]).

## § 4. Applications du principe du point fixe

Voyons quelques applications des théorèmes prouvés plus haut.

4.1. Soit donnée une fonction $\Phi(s, u)$ de deux variables réelles, définie sur la bande $a \leqslant s \leqslant b$, $-\infty < u < \infty$, on suppose que $\Phi$ est continue et possède une dérivée continue par rapport à $u$, telle que

$$0 < m \leqslant \Phi'_u(s, u) \leqslant M \quad (a \leqslant s \leqslant b, \quad -\infty < u < \infty). \tag{1}$$

Sous ces conditions, il existe une fonction unique $u = x^*(s)$ continue sur $[a, b]$ et vérifiant l'équation

$$\Phi(s, x^*(s)) = 0. \tag{2}$$

On peut établir ce résultat à l'aide du théorème 1.1. A cet effet, on considère dans l'espace $\mathbf{X} = \mathbf{C}[a, b]$ l'opérateur $P$:

$$y = P(x), \ y(s) = x(s) - \frac{2}{M+m}\Phi(s, x(s)) \quad (s \in [a, b]).$$

On vérifie aussitôt que $P$ est un opérateur de contraction. En effet, si $y = P(x)$, $\tilde{y} = P(\tilde{x})$, alors pour tout $s \in [a, b]$

$$|y(s) - \tilde{y}(s)| =$$
$$= \left| x(s) - \tilde{x}(s) - \frac{2}{M+m}[\Phi(s, x(s)) - \Phi(s, x(s))] \right| =$$
$$= |x(s) - \tilde{x}(s)| \left| 1 - \frac{2}{M+m}\Phi'_u(s, \theta(s)) \right| \leqslant \alpha \| x - \tilde{x} \|,$$

où

$$\alpha = \frac{M-m}{M+m},$$

donc $\alpha < 1$ en vertu de (1).

L'opérateur $P$ possède un point fixe unique $x^*$ dans $\mathbf{C}[a, b]$. Reste à remarquer que la relation $x^* = P(x^*)$ équivaut à l'égalité (2).

4.2. Considérons maintenant le système d'équations différentielles ordinaires

$$\frac{dx_i}{dt} = \varphi_i(x_1, x_2, \ldots, x_n; t) \quad (i = 1, 2, \ldots, n) \tag{3}$$

En écriture vectorielle le système (3) prend la forme de l'équation

$$\frac{dx}{dt} = \varphi(x, t), \tag{4}$$

où $x(t)$ est un vecteur de composantes $x_1(t), x_2(t), \ldots, x_n(t)$, et $\varphi(x, t)$ un vecteur de composantes $\varphi_1(x_1, \ldots, x_n; t), \ldots$ $\ldots, \varphi_n(x_1, x_2, \ldots, x_n; t)$. Ceci étant, par dérivée on entend la fonction vectorielle de composantes $\frac{dx_1}{dt}, \frac{dx_2}{dt}, \ldots, \frac{dx_n}{dt}$.

Supposons qu'on demande la solution du système (3), qui vérifie sur l'intervalle $[a, b]$ les conditions initiales

$$x_i(a) = y_0^{(i)} \quad (i = 1, 2, \ldots, n)$$

ou, sous la forme vectorielle,

$$x(a) = y_0 \tag{5}$$

($y_0$ est un élément de $\mathbf{R}^n$ de coordonnées $y_0^{(1)}, y_0^{(2)}, \ldots, y_0^{(n)}$). Résoudre ce problème revient à chercher la solution de l'équation intégrale

$$x(t) = y_0 + \int_a^t \varphi(x(\tau), \tau)\, d\tau. \tag{6}$$

THÉORÈME 1. *Soit $\varphi(y, t)$ une fonction vectorielle définie et bornée pour*

$$y \in \mathbf{R}^n, \quad |y - y_0| < \delta \ ^{*)}, \quad t \in [a, b'],$$

*mesurable relativement à t pour toute valeur de y et vérifiant la condition de Lipschitz par rapport à y:*

$$|\varphi(y, t) - \varphi(\tilde{y}, t)| \leqslant K\, |y - \tilde{y}|, \tag{7}$$

*où K ne dépend pas de t.*

*Soit*

$$M = \sup |\varphi(y, t)| \quad (|y - y_0| \leqslant \delta, \quad t \in [a, b']).$$

*Si l'intervalle $[a, b] \subset [a, b']$ est assez petit, plus exactement si*

$$b - a < \min\left[\frac{\delta}{M}, \frac{1}{K}\right], \tag{8}$$

*l'équation intégrale* (6), *donc le système d'équations différentielles* (3), *admet une solution unique vérifiant les conditions initiales données. De plus, la solution dépend continûment du vecteur initial $y_0$.*

DÉMONSTRATION. Considérons l'espace $\mathbf{C}_n = \mathbf{C}([a, b], \mathbf{R}^n)$ des fonctions vectorielles continues réelles de dimension $n$, définies sur $[a, b]$. $\mathbf{C}_n$, linéarisé de façon naturelle, devient un $B$-espace si l'on pose

$$\|x\| = \max_{t \in [a, b]} |x(t)|, \quad x \in \mathbf{C}_n.$$

---

*) Ici et dans la suite on désigne par $|y|$ la longueur euclidienne du vecteur $y$.

Dans le but d'appliquer le théorème 1.3 à l'équation (6), on pose $\mathbf{X} = \mathbf{C}_n$, $\mathbf{Y} = \mathbf{R}^n$. L'équation (6) peut s'écrire sous la forme de l'équation fonctionnelle

$$x = P_{y_0}(x), \tag{9}$$

où $P_y$ est un opérateur intégral :

$$\begin{gathered} z = P_y(x) \quad (y \in \mathbf{R}^n, \quad x \in \mathbf{C}_n,) \\ z(t) = y + \int_a^t \varphi(x(\tau), \tau)\, d\tau, \end{gathered} \tag{10}$$

défini sur l'ensemble $\Omega$ de tous les $x \in \mathbf{C}_n$ vérifiant

$$| x(t) - y_0 | \leqslant \delta \quad (t \in [a, b]).$$

Pour $\mathbf{Y}_0$ il convient de prendre l'ensemble des vecteurs $y \in \mathbf{R}^n$ tels que

$$| y - y_0 | < \delta_1 = \delta - M(b - a).$$

L'opérateur $P_y$ envoie $\Omega$ dans lui-même pour tout $y \in \mathbf{Y}_0$. En effet,

$$| z(t) - y_0 | \leqslant | y - y_0 | + \left| \int_a^t \varphi(x(\tau), \tau)\, d\tau \right| \leqslant \delta_1 + M(b-a) = \delta.$$

Vérifions que la condition (2) du § 1 est satisfaite. Soient $x$, $\tilde{x} \in \Omega$; $y \in \mathbf{Y}_0$,

$$\| P_y(x) - P_y(\tilde{x}) \| = \max_{t \in [a, b]} \left| \int_a^t [\varphi(x(\tau), \tau) - \varphi(\tilde{x}(\tau), \tau)]\, d\tau \right| \leqslant$$
$$\leqslant K(b-a) \| x - \tilde{x} \|.$$

En posant

$$\alpha = K(b - a)$$

on a $\alpha < 1$ en vertu de (8).

La continuité de l'opérateur $P_y$ en $y$ est évidente.

Donc, en vertu du théorème 1.3, l'équation (9) et, par suite, l'équation intégrale (6), admet une solution unique dépendant continûment de $y$.

Remarque. La dépendance continue de la solution $x_y^*$ du système (3) par rapport au vecteur initial $y$ pour $y = y_0$ exprime ici le fait que pour tout $\varepsilon > 0$ il existe $\eta > 0$, tel que $\| x_y^* - x_{y_0}^* \|_{\mathbf{C}} < \varepsilon$ dès que $| y - y_0 | < \eta$.

**4.3.** L'application du principe de Schauder à l'opérateur (9) permet d'établir l'existence de la solution de l'équation (6), donc du système (3), moyennant des conditions plus faibles sur la fonc-

tion $\varphi(y, t)$. Mais, dans ce cas, on ne peut démontrer l'unicité de la solution et *a fortiori* sa dépendance continue par rapport au vecteur initial.

THÉORÈME 2. *Supposons qu'une fonction* $\varphi$ *vérifie toutes les conditions du théorème* 1 *à l'exception de la condition* (7) *qui est remplacée par une plus faible, savoir: quel que soit* $t \in [a, b']$ *la fonction* $\varphi(y, t)$ *est continue en* $y$ *uniformément par rapport à* $t \in [a, b]$.

*Si*

$$b-a \leqslant \frac{\delta}{M},$$

*alors le système* (3) *admet au moins une solution vérifiant la condition initiale* (5).

DEMONSTRATION. On conserve les notations du théorème 1 et on pose $P = P_{y_0}$. Il est aisé de voir que l'opérateur $P$ applique $\Omega$ dans lui-même. En effet, en démontrant ce fait dans le théorème 1, on n'a utilisé que l'inégalité $M(b-a) \leqslant \delta$ qui est réalisée ici.

Il est immédiat aussi de vérifier que $P$ est un opérateur continu. Supposons en effet que $x_n \to x_0$, $x_n \in \Omega$, $z_n = P(x_n)$ $(n = 0, 1, \ldots)$. La fonction $\varphi(y, t)$ étant continue, pour $\varepsilon > 0$ il existe un $\eta > 0$ tel que $|y - y'| < \eta$ entraîne

$$|\varphi(y, t) - \varphi(y', t)| < \varepsilon \quad (t \in [a, b]). \tag{11}$$

Comme $\| x_n - x_0 \| \to 0$, pour $n$ assez grand $(n \geqslant n_0)$ on aura $\| x_n - x_0 \| < \eta$ et *a fortiori*

$$|x_n(t) - x_0(t)| < \eta \quad (t \in [a, b]).$$

En tenant compte de (11), on peut écrire pour les $n$ indiqués

$$|\varphi(x_n(t), t) - \varphi(x_0(t), t)| < \varepsilon \quad (t \in [a, b]).$$

Donc, pour $n \geqslant n_0$

$$\| z_n - z_0 \| = \max_{t \in [a, b]} |z_n(t) - z_0(t)| \leqslant$$

$$\leqslant \max_{t \in [a, b]} \int_a^t |\varphi(x_n(\tau), \tau) - \varphi(x_0(\tau), \tau)\, d\tau \leqslant \varepsilon (b-a).$$

Donc

$$z_n = P(x_n) \to z_0 = P(x_0).$$

Il est clair que $\Omega$ est un ensemble fermé convexe, donc pour appliquer le théorème 3.2, il faut établir seulement la compacité relative de l'ensemble $P(\Omega)$. Ce qui achèvera la démonstration du théorème.

Le critère de compacité d'un ensemble dans l'espace **C** (cf. I.5.4, tome 1) se généralise sans peine à l'espace $\mathbf{C}_n$. Il faut donc prouver que les fonctions de la famille $P(\Omega)$ sont équicontinues (ces fonc-

tions sont uniformément bornées, puisque $P(\Omega) \subset \Omega$ et $\Omega$ est borné). Or, ceci est une conséquence de l'inégalité

$$|z(t') - z(t)| = \left| \int_t^{t'} \varphi(x(\tau), \tau)\, d\tau \right| \leqslant M |t' - t| \quad (x \in \Omega, \ z = P(x)).$$

C. q. f. d.

4.4. Considérons l'équation intégrale

$$x(s) = \lambda \int_a^b \varphi(s, t, x(t))\, dt. \tag{12}$$

En se servant du théorème 1.1, on prouve que l'équation (12) admet une solution unique pour $\lambda$ assez petit.

THÉORÈME 3. *Soit* $\varphi(s, t, u)$ *une fonction définie et continue dans le parallélépipède*

$$a \leqslant s, \quad t \leqslant b, \quad |u| \leqslant h$$

*et vérifiant* ($\forall s$ *et* $\forall t$ *fixes*) *la condition de Lipschitz par rapport à* $u$:

$$|\varphi(s, t, u) - \varphi(s, t, u')| \leqslant K |u - u'|,$$

$K$ *ne dépendant pas de* $s$ *et de* $t$.

*Si l'on désigne*

$$M = \max |\varphi(s, t, u)| \quad (a \leqslant s, \quad t \leqslant b, \quad |u| \leqslant h),$$

*alors l'équation* (12) *admet une solution unique pour les* $\lambda$ *tels que*

$$|\lambda| < \min \left[ \frac{h}{M(b-a)}, \ \frac{1}{K(b-a)} \right]. \tag{13}$$

DEMONSTRATION. Posons $\mathbf{X} = \mathbf{C}[a, b]$ dans le théorème 1.1 et prenons pour $\Omega$ l'ensemble des $x \in \mathbf{C}$ tels que $\| x \| \leqslant h$. Définissons l'opérateur $P$ comme

$$z = P(x), \ z(s) = \lambda \int_a^b \varphi(s, t, x(t))\, dt.$$

La condition (13) dit que $P(x) \in \Omega$. En effet,

$$\| P(x) \| \leqslant |\lambda| M (b-a) < h \quad (x \in \Omega).$$

D'autre part

$$\| P(x) - P(x') \| \leqslant |\lambda| \max_{s \in [a, b]} \int_a^b |\varphi(s, t, x(t)) -$$
$$- \varphi(s, t, x'(t))|\, dt \leqslant |\lambda| K (b-a) \| x - x' \|.$$

Si donc l'on pose $\alpha = |\lambda| K(b-a)$, on déduit de (13) que $\alpha < 1$, c'est-à-dire $P$ est un opérateur de contraction.

L'équation (12) pouvant être mise sous la forme

$$x = P(x), \tag{14}$$

l'existence et l'unicité de la solution de l'équation (12) résultent du théorème 1.1 dont toutes les conditions, on vient de le voir, sont satisfaites.

Remarque 1. Prenons une fonction continue $x_0$ ($\|x_0\| \leqslant h$) et formons à partir d'elle la suite $\{x_n\}$ de fonctions

$$x_{n+1}(s) = \lambda \int_a^b \varphi(s, t, x_n(t))\, dt \quad (n = 0, 1, \ldots).$$

Le théorème 1.1 nous dit que cette suite est uniformément convergente vers la solution de l'équation (12).

Remarque 2. En se servant du théorème 1.3 on peut montrer que la solution de l'équation (12) dépend continûment de $\lambda$, pour les $\lambda$ vérifiant (13).

4.5. L'utilisation du principe de Schauder permet parfois d'élargir le domaine des $\lambda$ pour lesquelles l'équation (12) admet une solution.

Supposons que la fonction $\varphi$ est de la forme

$$\varphi(s, t, u) = K(s, t)\, \psi(t, u), \tag{15}$$

où $K(s, t)$ est une fonction continue dans le carré $[a, b; a, b]$ et $\psi(t, u)$ continue pour $t \in [a, b]$, $|u| \leqslant h$. On a le

Théorème 4. *Si dans l'équation intégrale* (12), *la fonction* $\varphi$ *est de la forme* (15) *et*

$$|\lambda| \leqslant \frac{h}{M_1 M_2 (b-a)}, \tag{16}$$

*où* $M_1 = \max\limits_{a \leqslant b,\, t \leqslant b} |K(s, t)|$, $M_2 = \max\limits_{\substack{a \leqslant t \leqslant b \\ |u| \leqslant h}} |\psi(t, u)|$, *alors l'équation* (12) *admet une solution continue.*

Démonstration. Conservons les notations précédentes et prouvons comme plus haut que l'opérateur $P$ envoie l'ensemble $\Omega$ dans lui-même. L'ensemble $\Omega$ étant convexe et fermé, et l'opérateur $P$ continu (cf. démonstration du théorème 2), il suffit comme dans le théorème 2 de vérifier que $P(\Omega)$ est relativement compact; on pourra alors utiliser le théorème 3.2. L'ensemble $P(\Omega)$ étant un sous-ensemble de $C[a, b]$, il faut prouver que les fonctions de $P(\Omega)$ sont 1) uniformément bornées et 2) équicontinues. 1) est évident.

2) est une conséquence de l'inégalité ($z = P(x)$, $x \in \Omega$)

$$|z(s) - z(s')| = |\lambda| \left| \int_a^b [K(s, t) - K(s', t)] \psi(t, x(t))\, dt \right| \leqslant$$

$$\leqslant |\lambda| M_2 \int_a^b |K(s, t) - K(s', t)|\, dt.$$

En traitant l'opérateur intégral comme un opérateur dans l'espace $\mathbf{L}^2[a, b]$, on peut établir grâce au principe de Schauder un résultat analogue au théorème 4.

THÉORÈME 5. *Supposons que dans l'équation* (12) *la fonction* $\varphi(s, t, u)$ *est de la forme* (15) *et de plus que la fonction* $K(s, t)$ *est mesurable dans le carré* $[a, b; a, b]$ *et vérifie la condition*

$$A^2 = \sup_{s \in [a, b]} \int_a^b |K(s, t)|^2\, dt < \infty,$$

*et la fonction* $\psi(t, u)$ *définie et continue dans le rectangle* $[a, b; -h, h]$.
*Si*

$$|\lambda| \leqslant \frac{h}{AM_2 \sqrt{b-a}}, \tag{7}$$

*alors l'équation* (12) *admet une solution de carré sommable sur* $[a, b]$.

DÉMONSTRATION. Comme plus haut on fait la démonstration en se servant du théorème 3.2. Mais on pose ici $\mathbf{X} = \mathbf{L}^2[a, b]$, on prend pour $\Omega$ l'ensemble des $x \in \mathbf{L}^2$ tels que

$$|x(t)| \leqslant h \quad \text{pour presque tous les } t \in [a, b],$$

et l'on conserve le même $P$.

Vérifions que la condition (17) assure que $P(\Omega) \subset \Omega$. En effet, si $x \in \Omega$ et $z = P(x)$, alors pour $s \in [a, b]$ l'inégalité de Bouniakovski nous dit que

$$|z(s)| = |\lambda| \left| \int_a^b K(s, t) \psi(t, x(t))\, dt \right| \leqslant$$

$$\leqslant |\lambda| \left[ \int_a^b |K(s, t)|^2 \right]^{1/2} \left[ \int_a^b |\psi(t, x(t))|^2\, dt \right]^{1/2} \leqslant$$

$$\leqslant |\lambda| A M_2 \sqrt{b-a} \leqslant h.$$

Prouvons que l'opérateur $P$ est continu. Soient $x, x' \in \Omega$ et $z = P(x)$, $z' = P(x')$. On a

$$|z(s) - z'(s)|^2 \leqslant$$
$$\leqslant |\lambda|^2 \int_a^b |K(s, t)|^2 dt \int_a^b |\psi(t, x(t)) - \psi(t, x'(t))|^2 dt \leqslant$$
$$\leqslant |\lambda|^2 A^2 \int_a^b |\psi(t, x(t)) - \psi(t, x'(t))|^2 dt. \quad (18)$$

Soit donné un $\varepsilon > 0$. La fonction $\psi$ étant continue, il existe $\eta > 0$ tel que pour $|u - u'| < \eta$ l'on ait

$$|\psi(t, u) - \psi(t, u')| < \varepsilon \quad (t \in [a, b]).$$

Partageons l'intervalle $[a, b]$ en deux ensembles $e_1$ et $e_2$, en rapportant à $e_1$ les points en lesquels la différence $|x(t) - x'(t)| < \eta$, et à $e_2$, les autres. En posant $\|x - x'\| = \gamma$, on obtient

$$\gamma^2 = \int_a^b |x(t) - x'(t)|^2 dt \geqslant \int_{e_2} |x(t) - x'(t)|^2 dt \geqslant \eta^2 \operatorname{mes} e_2,$$

de sorte que

$$\operatorname{mes} e_2 \leqslant \gamma^2/\eta^2.$$

Majorons maintenant l'intégrale $I = \int_a^b |\psi(t, x(t)) - \psi(t, x(t))|^2 dt$:

$$I = \int_{e_1} + \int_{e_2} |\psi(t, x(t)) - \psi(t, x'(t))|^2 dt \leqslant$$
$$\leqslant \varepsilon^2 \operatorname{mes} e_1 + 4M_2^2 \operatorname{mes} e_2 \leqslant \varepsilon^2 (b-a) + \frac{4M_2^2\gamma^2}{\eta^2}.$$

En utilisant cette majoration, on obtient à partir de (18)

$$\|z - z'\|^2 = \int_a^b |z(s) - z'(s)|^2 ds \leqslant$$
$$\leqslant |\lambda|^2 A^2 (b-a) \left[\varepsilon^2 (b-a) + \frac{4M_2^2 \|x - x'\|^2}{\eta^2}\right].$$

Donc $z$ et $z'$ peuvent être rendus aussi proches que l'on veut si $x$ est assez proche de $x'$.

L'ensemble $\Omega$ étant visiblement convexe et fermé, il reste à prouver la compacité relative de l'ensemble $P(\Omega)$. A cet effet, utilisons le théorème de M. Riesz (IX.1.3, tome 1) qui dit qu'il suffit de véri-

fier que

$$\int_a^b |z(t+\eta)-z(t)|^2\,dt \xrightarrow[\eta\to 0]{} 0$$

uniformément en $z \in P(\Omega)$ *).

Soient $x \in \Omega$, $z = P(x)$. On a alors

$$\int_a^b |z(s+\eta)-z(s)|^2\,ds =$$

$$= \int_a^b \left|\lambda \int_a^b [K(s+\eta,\ t)-K(s,\ t)]\,\psi(t,\ x(t))\,dt\right|^2 ds \leqslant$$

$$\leqslant M_2^2|\lambda|^2 \int_a^b \int_a^b |K(s+\eta,\ t)-K(s,\ t)|^2\,ds\,dt.$$

Montrons que la dernière intégrale tend vers zéro.

Ceci est évident si la fonction $K(s,\ t)$ est continue. Si $K(s,\ t)$ est générique, on peut exhiber une suite $\{K_n(s,\ t)\}$ de fonctions continues, telle que

$$\int_a^b \int_a^b |K(s,\ t)-K_n(s,\ t)|^2\,ds\,dt \xrightarrow[n\to\infty]{} 0.$$

On a

$$\left[\int_a^b \int_a^b |K(s+\eta,\ t)-K(s,\ t)|^2\,ds\,dt\right]^{1/2} \leqslant$$

$$\leqslant \left[\int_a^b \int_a^b |K_n(s+\eta,\ t)-K_n(s,\ t)|^2\,ds\,dt\right]^{1/2} +$$

$$+ \left[\int_a^b \int_a^b |K(s+\eta,\ t)-K_n(s+\eta,\ t)|^2\,ds\,dt\right]^{1/2} +$$

$$+ \left[\int_a^b \int_a^b |K(s,\ t)-K_n(s,\ t)|^2\,ds\,dt\right]^{1/2}.$$

En choisissant $n$ assez grand pour que les deux dernières intégrales du second membre soient $<\varepsilon$ ($\varepsilon > 0$), et ensuite $\eta$ assez petit pour

*) Le théorème de M. Riesz comporte encore une condition indiquant que les normes sont bornées. Cette condition n'est pas mentionnée dans le texte, car elle est évidente.

que la première intégrale soit également $<\varepsilon$, on obtient le résultat souhaité dans le cas général.

C.q.f.d.

REMARQUE 1. Si la fonction $\psi(t, u)$ est définie dans la bande $a \leqslant t \leqslant b$, $-\infty < u < \infty$ et si de plus

$$\frac{\psi(t, h)}{h} \xrightarrow[h\to\infty]{} 0, \tag{19}$$

alors pour $h$ assez grand, la condition (16) du théorème 4 et la condition (17) du théorème 5 seront satisfaites quels que soient $\lambda$, $M_1$, $M_2$ et $b - a$. Donc, (19) assure l'existence de la solution de (12) pour tout $\lambda$.

REMARQUE 2. Les théorèmes 4 et 5 peuvent être démontrés sous d'autres conditions portant sur les fonctions figurant dans les équations. Il est de même aisé de formuler des théorèmes dans lesquels l'équation (12) serait traitée dans un espace fonctionnel différent de $C[a, b]$ et de $L^2(a, b)$ (par exemple dans $L^1(a, b)$).

Pour l'application du principe du point fixe à la démonstration de l'existence d'une solution des équations non linéaires voir Krasnosselski [1].

D'importantes applications du principe de Schauder sont données dans ses travaux sur les équations différentielles elliptiques. Des applications des théorèmes de Leray-Schauder sont accessibles dans les travaux sur les équations hyperboliques (cf. Schauder [3]).

## § 5. Théorème de Kakutani

5.1. Le théorème de Kakutani (cf. Kakutani [2]) est une généralisation du théorème de Schauder (3.1). Le théorème de Kakutani connaît d'innombrables applications en économie mathématique (cf. Nikaïdo). Plus bas on indique une application de ce théorème à la théorie des jeux.

Pour énoncer le théorème de Kakutani nous aurons besoin de quelques faits sur les applications multivalentes. Soient $X$ et $Y$ des espaces métriques; on désigne par $2^Y$ l'ensemble de toutes les parties de $Y$. Toute application $f: X \to 2^Y$ sera dite *application multivalente*, c'est-à-dire l'application $f$ associe à tout point $x \in X$ un sous-ensemble $f(x)$ de $Y$, que nous supposerons être non vide.

On appelle *graphe* de l'application multivalente $f: X \to 2^Y$ le sous-ensemble $G_f$ de $X \times Y$:

$$G_f = \{(x, y): y \in f(x)\}.$$

L'application $f: X \to 2^Y$ est *fermée en* $x \in X$ si la convergence des suites $x_n \to x$, $y_n \to y$ et $y_n \in f(x_n)$ entraînent $y \in f(x)$. L'application $f$ est *fermée* si elle l'est en chaque point $x \in X$. Il est manifeste que $f$ est fermée si et seulement si son graphe $G_f$ l'est dans $X \times Y$.

L'application multivalente $f: X \to 2^Y$ est *semi-continue supérieurement en* $x \in X$ si pour tout ensemble ouvert $U$ tel que $f(x) \subset U$

il existe un voisinage $V$ de $x$ tel que $f(V) \subset U$, où $f(V) = \bigcup_{z \in V} f(z)$. L'application $f$ est dite *semi-continue supérieurement* si elle l'est en chaque point $x \in X$.

Si $f$ est une application univalente ordinaire et si elle est semi-continue supérieurement, elle est continue, tandis qu'une application fermée peut ne pas l'être.

LEMME 1. *Si l'ensemble* Y *est compact, l'application fermée $f$ est semi-continue supérieurement.*

DÉMONSTRATION. Soient $x \in X$ et un ensemble ouvert $U \supset f(x)$. Considérons l'ensemble $V = \{z \in X : f(z) \subset U\}$. Comme $x \in V$ et $f(V) \subset U$, il nous suffit de montrer que $V$ est ouvert ou bien que $X \setminus V = \{z \in X : f(z) \cap (Y \setminus U) \neq \emptyset\}$ est fermé. Si la suite $z_n \to z \in X$ et $z_n \in X \setminus V$, il existe des points $y_n \in f(z_n) \cap (Y \setminus U)$. L'ensemble Y étant compact, il existe une suite partielle $\{y_{n_k}\}$ telle que $y_{n_k} \to y \in Y \setminus U$. L'application $f$ étant fermée, il s'ensuit que $y \in f(z)$. Donc, $y \in f(z) \cap (Y \setminus U)$, d'où $z \in X \setminus V$.

5.2. Un point $x^* \in X$ est *point fixe* de l'application $f: X \to 2^X$ si $x^* \in f(x^*)$. Nous sommes à présent en mesure d'énoncer le théorème de Kakutani.

THÉORÈME 1. *Soient* X *un B-espace, $K$ un ensemble convexe compact non vide dans* X, *et $f: K \to 2^K$ une application multivalente satisfaisant aux conditions*:

1) *pour tout point $x \in K$ l'ensemble $f(x)$ est un sous-ensemble convexe non vide de $K$;*

2) *l'application $f$ est fermée.*

*Alors l'application $f$ possède un point fixe.*

DÉMONSTRATION *). L'ensemble $K$ étant compact, le théorème de Hausdorff affirme l'existence pour tout $\varepsilon > 0$ d'un $\varepsilon$-réseau fini pour $K$: $x_{\varepsilon 1}, x_{\varepsilon 2}, \ldots, x_{\varepsilon n(\varepsilon)}$. Pour tout $x \in K$ posons

$$\varphi_{\varepsilon i}(x) = \max(\varepsilon - \| x - x_{\varepsilon i} \|, 0) \quad (i = 1, \ldots, n(\varepsilon)).$$

Il est évident que $\varphi_{\varepsilon i}$ est une fonction continue positive sur $K$. Comme $\{x_{\varepsilon i}\}_{i=1}^{n(\varepsilon)}$ est un $\varepsilon$-réseau, pour tout $x \in K$ on a $\| x - x_{\varepsilon i} \| < \varepsilon$ pour au moins un $i$. Donc, pour cet $i$ on a $\varphi_{\varepsilon i}(x) > 0$. Ce raisonnement prouve la validité de définition suivante:

$$w_{\varepsilon i}(x) = \frac{\varphi_{\varepsilon i}(x)}{\sum_{j=1}^{n(\varepsilon)} \varphi_{\varepsilon j}(x)} \quad (i = 1, \ldots, n(\varepsilon)).$$

Figeons maintenant un point quelconque $y_{\varepsilon i} \in f(x_{\varepsilon i})$ ($i = 1, \ldots, n(\varepsilon)$) et définissons une application continue univalente

*) Nous suivons Nikaïdo pour démontrer le théorème 1.

$f_\varepsilon : K \to K$ par la formule

$$f_\varepsilon(x) = \sum_{i=1}^{n(\varepsilon)} w_{\varepsilon i}(x)\, y_{\varepsilon i}.$$

Les conditions $y_{\varepsilon i} \in K$, $w_{\varepsilon i}(x) \geqslant 0$, $\sum_i w_{\varepsilon i}(x) = 1$ et la convexité de $K$ entraînent que $f_\varepsilon(x) \in K$. Donc, pour tout $\varepsilon > 0$ est définie l'application continue univalente $f_\varepsilon : K \to K$. D'après le théorème de Schauder (3.1), cette application possède un point fixe $x_\varepsilon$: $f_\varepsilon(x_\varepsilon) = x_\varepsilon$.

L'ensemble $K$ étant compact, il existe une suite $\{\varepsilon_n\}$ de nombres strictement positifs et un point $x^* \in K$ tels que:

a) $\lim\limits_{n\to\infty} \varepsilon_n = 0$;

b) $x_{\varepsilon_n} \to x^*$;

c) $f_{\varepsilon_n}(x_{\varepsilon_n}) = x_n$.

Prouvons que $x^*$ est le point fixe de l'application $f$ cherché. Posons $U_\delta = f(x^*) + B_\delta$, où $B_\delta = \{y : \|y\| < \delta\}$, $\delta > 0$. Nous allons montrer que $x^* \in U_\delta$, $\forall \delta > 0$, d'où, puisque $f(x^*)$ est fermé, l'on obtiendra $x^* \in f(x^*)$ (la fermeture de $f(x^*)$ découle de celle de $f$).

Il est évident que $U_\delta$ est un ensemble ouvert convexe et $f(x^*) \subset U_\delta$. Le lemme 1 nous dit qu'on peut trouver pour $x^*$ et $U_\delta$ une boule $K_\varepsilon = \{x : \|x - x^*\| < \varepsilon\}$ telle que $f(K_\varepsilon) \subset U_\delta$. D'après les conditions a) et b) il existe un nombre $N$ tel que pour $n \geqslant N$ on a $\varepsilon_n < \varepsilon/2$ et $x_{\varepsilon_n} \in K_{\varepsilon/2}$. Si $w_{\varepsilon_n i}(x_{\varepsilon_n}) > 0$, alors

$$\| x_{\varepsilon_n i} - x_{\varepsilon_n} \| < \varepsilon_n < \varepsilon/2$$

et

$$\| x_{\varepsilon_n i} - x^* \| \leqslant \| x_{\varepsilon_n i} - x_{\varepsilon_n} \| + \| x_{\varepsilon_n} - x^* \| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Donc, pour $n > N$ on a $x_{\varepsilon_n i} \in K_\varepsilon$ pour tous les $i$, tels que $w_{\varepsilon_n i}(x_{\varepsilon_n}) > 0$. Pour ces $i$ on obtient

$$y_{\varepsilon_n i} \in f(x_{\varepsilon_n i}) \subset f(K_\varepsilon) \subset U_\delta. \tag{1}$$

De la condition c) on déduit

$$x_{\varepsilon_n} = \sum w_{\varepsilon_n i}(x_{\varepsilon_n})\, y_{\varepsilon_n i}. \tag{2}$$

Les relations (1) et (2) entraînent que pour $n \geqslant N$ le point $x_{\varepsilon_n}$ est combinaison convexe seulement des points $y_{\varepsilon_n i}$ appartenant à $U_\delta$, d'où $x_{\varepsilon_n} \in U_\delta$, puisque $U_\delta$ est convexe. En passant à la limite pour $n \to \infty$ on obtient $x^* \in \overline{U_\delta} \subset U_{2\delta}$. Mais on a déjà vu que ceci implique que $x^* \in f(x^*)$. Ceci achève la démonstration du théorème de Kakutani.

5.3. Voici l'application du théorème de Kakutani à la théorie des jeux.

Nous considérons un jeu à deux joueurs (I et II) de somme nulle, c'est-à-dire qu'un joueur gagne ce que l'autre perd. On appelle *stratégie* d'un joueur la description complète de toutes les actions qu'il a l'intention d'entreprendre dans chaque situation qui se présentera au cours du jeu. On désignera par $X$ l'ensemble de toutes les stratégies (on dit encore espace des stratégies) du joueur I et par $Y$ celui du joueur II.

On appelle *fonction de gain* une fonction réelle $K(x, y)$ définie sur $X \times Y$, le nombre $K(x, y)$ étant traité comme le gain du joueur I s'il adopte la stratégie $x \in X$ et le joueur II la stratégie $y \in Y$. Le nombre $-K(x, y)$ est interprété comme le gain du joueur II dans la même situation. On appelle *jeu* le triplet $(X, Y, K)$, où $X$ et $Y$ sont des ensembles, $K(x, y)$ une fonction sur $X \times Y$.

Lorsque le joueur I choisit une stratégie $x \in X$ sans rien savoir de la stratégie du joueur II, il doit admettre que son adversaire choisit sa meilleure stratégie, c'est-à-dire celle qui minimisera le gain du joueur I, soit $\inf_{y \in Y} K(x, y)$. Donc, le joueur I doit chercher une stratégie $x_0 \in X$ telle que

$$\inf_{y \in Y} K(x_0, y) = \max_{x \in X} \inf_{y \in Y} K(x, y). \tag{3}$$

Le gain du joueur II étant égal à $-K(x, y)$, de (3) l'on déduit que II doit chercher une stratégie $y_0 \in Y$ telle que

$$\inf_{x \in X} \{-K(x, y_0)\} = \max_{y \in Y} \inf_{x \in X} [-K(x, y)],$$

ou, ce qui revient au même,

$$\sup_{x \in X} K(x, y_0) = \min_{y \in Y} \sup_{x \in X} K(x, y). \tag{4}$$

La comparaison de (3) et (4) montre qu'une condition nécessaire et suffisante pour qu'existe des stratégies $x_0 \in X$ et $y_0 \in Y$ vérifiant (3) et (4) est que pour tous les $x \in X$ et $y \in Y$ l'on ait

$$K(x, y_0) \leqslant K(x_0, y_0) \leqslant K(x_0, y). \tag{5}$$

Pour que ceci ait lieu il faut et il suffit que

$$\min_{y \in Y} \sup_{x \in X} K(x, y) = \max_{x \in X} \inf_{y \in Y} K(x, y). \tag{6}$$

La valeur commune des deux membres de (6) est appelée *valeur* du jeu. Tout couple de stratégies $x_0$, $y_0$ vérifiant les inégalités (5) est appelé *solution* du jeu, et les stratégies $x_0$, $y_0$ *stratégies optimales* *). Pour prouver l'existence d'une solution du jeu il faut prouver

*) Pour plus de détails voir l'ouvrage de Karlin dont nous nous inspirons parfois dans notre exposé.

le théorème de *minimax* qui consiste en la vérification de l'égalité (6). Nous allons nous servir du théorème de Kakutani pour prouver un théorème très général de cette nature. Nous énonçons tout d'abord un lemme dont nous laissons la démonstration au lecteur.

LEMME 2. 1) *Pour toute fonction réelle* $K(x, y)$ *on a*

$$\inf_{y \in Y} \sup_{x \in X} K(x, y) \geqslant \sup_{x \in X} \inf_{y \in Y} K(x, y). \tag{7}$$

2) *Si* $X$ *et* $Y$ *sont des compacts métriques et la fonction* $K(x, y)$ *continue sur* $X \times Y$, *alors les fonctions* $\varphi(x) = \max_{y \in Y} K(x; y)$ *et* $\psi(x) = \min_{y \in Y} K(x, y)$ *sont continues sur* $X$. *De plus existent* $\min_{y \in Y} \max_{x \in X} K(x, y)$ *et* $\max_{x \in X} \min_{y \in Y} K(x, y)$.

THÉORÈME 2. *Soient réalisées les conditions suivantes:*
a) $X$ *est un ensemble compact convexe dans un B-espace* $\mathbf{X}$;
b) $Y$ *est un ensemble compact convexe dans un B-espace* $\mathbf{Y}$;
c) *la fonction* $K(x, y)$ *est continue sur* $X \times Y$, *convexe en* $y$ *pour tout* $x \in X$, *concave en* $x$ *pour tout* $y \in Y$, *c'est-à-dire*

$$K(x, \lambda y_1 + (1 - \lambda) y_2) \leqslant \lambda K(x, y_1) + (1 - \lambda) K(x_2, y),$$

$$K(\lambda x_1 + (1 - \lambda) x_2, y) \geqslant \lambda K(x_1, y) + (1 - \lambda) K(x_2, y),$$

*où* $x, x_1, x_2 \in X$; $y, y_1, y_2 \in Y$; $0 \leqslant \lambda \leqslant 1$.
*Alors le jeu* $(X, Y, K)$ *admet une solution et de plus*

$$\min_{y \in Y} \max_{x \in X} K(x, y) = \max_{x \in X} \min_{y \in Y} K(x, y). \tag{8}$$

DÉMONSTRATION. D'après 2) du lemme 2, la condition (6), qui exprime le fait que le jeu admet une solution, est confondue avec (8). Donc, c'est (8) qu'on prouvera.

Pour tout $x \in X$ posons

$$B_x = \{y \in Y : K(x, y) = \min_{z \in Y} K(x, z)\}.$$

L'ensemble $Y$ étant compact, on a $B_x \neq \varnothing$. Il est évident que $B_x$ est fermé. Montrons qu'il est convexe. Si $y_1, y_2 \in B_x$, $0 \leqslant \lambda \leqslant 1$, alors

$$K(x, \lambda y_1 + (1 - \lambda) y_2) \leqslant \lambda K(x, y_1) + (1 - \lambda) K(x, y_2) =$$
$$= \lambda \min_{z \in Y} K(x, z) + (1 - \lambda) \min_{z \in Y} K(x, z) = \min_{z \in Y} K(x, z). \tag{9}$$

D'autre part, $Y$ étant convexe, on a $\lambda y_1 + (1 - \lambda) y_2 \in Y$, d'où

$$K(x, \lambda y_1 + (1 - \lambda) y_2) \geqslant \min_{z \in Y} K(x, z). \tag{10}$$

13—01292

En combinant (9) et (10), on trouve

$$K(x, \lambda y_1 + (1-\lambda) y_2) = \min_{z \in Y} K(x, z).$$

Donc, $\lambda y_1 + (1-\lambda) y_2 \in B_x$, c'est-à-dire l'ensemble $B_x$ est convexe.

De façon analogue, pour tout $y \in Y$, l'ensemble

$$A_y = \{x \in X : K(x, y) = \max_{w \in X} K(w, y)\}$$

n'est pas vide, est fermé et convexe.

Soit un ensemble convexe $C = X \times Y$ dans un $B$-espace $\mathbf{X} \times \mathbf{Y}$. D'après le théorème de Tikhonov (I.2.8, tome 1) l'ensemble $C$ est compact dans $\mathbf{X} \times \mathbf{Y}$. Il est évident que $A_y \times B_x$ ($x \in X$, $y \in Y$) sont des sous-ensembles non vides, convexes, fermés dans $C$. Soit maintenant l'application multivalente $f: C \to 2^C$ qui associe à tout point $(x, y) \in C$ l'ensemble $A_y \times B_x$. Pour que l'on puisse user du théorème de Kakutani il nous reste à prouver que l'application $f$ est fermée.

Soient $(x_n, y_n) \to (x, y)$, $(u_n, v_n) \to (u, v)$ dans $C$ et $(u_n, v_n) \in f(x_n, y_n)$. Montrons que $(u, v) \in f(x, y) = A_y \times B_x$. Comme $u_n \in A_{y_n}$, $v_n \in B_{x_n}$, on a

$$K(u_n, y_n) = \max_{w \in X} K(w, y_n)$$

et

$$K(x_n, v_n) = \min_{z \in Y} K(x_n, z).$$

Comme $u_n \to u$, en vertu de 2) du lemme 2 on a

$$K(u, y) = \lim_{n\to\infty} K(u_n, y_n) = \lim_{n\to\infty} \max_{w \in X} K(w, y_n) = \max_{w \in X} K(w, y).$$

Donc, $u \in A_y$. De façon analogue, $v \in B_x$.

Par suite, l'application $f$ est fermée et l'on peut utiliser le théorème de Kakutani qui affirme l'existence d'un point $(x_0, y_0) \in C$, tel que $(x_0, y_0) \in f(x_0, y_0) = A_{y_0} \times B_{x_0}$. Donc

$$K(x_0, y_0) = \max_{x \in X} K(x, y_0) \quad \text{et} \quad K(x_0, y_0) = \min_{y \in Y} K(x_0, y).$$

D'où

$$\min_{y \in Y} \max_{x \in X} K(x, y) \leqslant K(x_0, y_0) \leqslant \max_{x \in X} \min_{y \in Y} K(x, y).$$

En combinant cette inégalité avec (7), on obtient (8). C.q.f.d.

CHAPITRE XVII

# DÉRIVATION DES OPÉRATEURS NON LINÉAIRES

L'étude ultérieure des opérateurs non linéaires passe par l'établissement de relations avec les opérateurs linéaires, plus exactement par une approximation locale des opérateurs non linéaires par des opérateurs linéaires.

Cette étude implique un calcul différentiel pour opérateurs non linéaires qui sera développé dans le cadre de ce chapitre mais appliqué seulement dans le chapitre suivant.

## § 1. Dérivée première

**1.1.** Soit donné un opérateur $P$ d'un ouvert $\Omega$ d'un $B$-espace **X** dans un ensemble $\Delta$ d'un autre $B$-espace **Y**. Fixons un élément $x_0 \in \Omega$ et supposons qu'il existe un opérateur linéaire continu $U \in B(\mathbf{X}, \mathbf{Y})$ *) tel que pour tout $x \in \mathbf{X}$

$$\lim_{t \to 0} \frac{P(x_0 + tx) - P(x_0)}{t} = U(x). \tag{1}$$

On dit que l'opérateur linéaire $U$ est la *dérivée* de l'opérateur $P$ au point $x_0$, et l'on note

$$U = P'(x_0).$$

Cette dérivée est souvent appelée *dérivée de Gateaux* ou *dérivée faible*, et l'élément $U(x)$, *différentielle de Gateaux*.

Désignons par $\overline{K}$ l'ensemble des $x \in \mathbf{X}$ tels que $\|x\| = 1$. Si la relation (1) est réalisée uniformément en $x \in \overline{K}$, on dit que l'opérateur $P$ est *dérivable* au point $x_0$ et la dérivée $P'(x_0)$ est appelée dans ce cas *dérivée de Fréchet* ou *dérivée forte* **).

La dérivabilité de l'opérateur $P$ au point $x_0$ signifie, en d'autres termes, qu'il existe un opérateur linéaire $U \in B(\mathbf{X}, \mathbf{Y})$ tel que pour

*) On rappelle que $B(\mathbf{X}, \mathbf{Y})$ désigne l'espace de tous les opérateurs linéaires continus de **X** dans **Y**.

**) La définition de la dérivée de Fréchet figure dans le travail de Fréchet [3], celle de la dérivée forte dans Gateaux [1].

tout $\varepsilon > 0$ on peut exhiber $\delta > 0$ tel que $\| \Delta x \| < \delta$ ($\Delta x \in X$) entraîne

$$\| P(x_0 + \Delta x) - P(x_0) - U(\Delta x) \| \leqslant \varepsilon \| \Delta x \|. \tag{2}$$

La démonstration simple de cette proposition est laissée au lecteur.

**1.2.** Indiquons quelques propriétés simples de la dérivée.

I. Si l'opérateur $P$ est dérivable en un point $x_0$, il est continu en ce point *).

Ceci résulte directement de (2).

II. Si $P = \alpha_1 P_1 + \alpha_2 P_2$, l'existence des dérivées $P_1'(x_0)$ et $P_2'(x_0)$ entraîne celle de

$$P'(x_0) = \alpha_1 P_1'(x_0) + \alpha_2 P_2'(x_0).$$

En outre, la dérivabilité de $P_1$ et de $P_2$ en $x_0$ entraîne celle de $P$.

C'est une conséquence immédiate de la définition.

III. Si $P = U \in B(X, Y)$, $P$ est dérivable en chaque point $x_0 \in X$ et

$$P'(x_0) = U.$$

En effet

$$\frac{P(x_0 + tx) - P(x_0)}{t} = U(x).$$

IV. Soient $P$ un opérateur de $\Omega \subset X$ dans un ouvert $\Delta \subset Y$, $Q$ un opérateur de $\Delta$ dans un $B$-espace Z. Posons $R = QP$ et supposons que $Q$ est dérivable au point $y_0 = P(x_0)$ ($x_0 \in \Omega$) et $P$ au point $x_0$. L'opérateur $R$ admet alors une dérivée en $x_0$ et

$$R'(x_0) = Q'(P(x_0))\, P'(x_0) = Q'(y_0)\, P'(x_0).$$

En effet, appelons $U = P'(x_0)$, $V = Q'(y_0)$ et prenons un $x \in X$. Posons $\Delta y = P(x_0 + tx) - P(x_0)$. Alors

$$\frac{R(x_0+tx) - R(x_0)}{t} = \frac{QP(x_0+tx) - QP(x_0)}{t} = \frac{Q(y_0+\Delta y) - Q(y_0)}{t} =$$
$$= \frac{V(\Delta y) + \zeta(\Delta y)}{t} = V\left(\frac{P(x_0+tx) - P(x_0)}{t}\right) +$$
$$+ \frac{\zeta(\Delta y)}{\| \Delta y \|} \left\| \frac{P(x_0+tx) - P(x_0)}{t} \right\|, \tag{3}$$

où $\zeta(\Delta y) = Q(y_0 + \Delta y) - Q(y_0) - V(\Delta y)$, $\| \zeta(\Delta y) \| = o(\| \Delta y \|)$. Pour $t \to 0$, le premier terme de (3) a pour limite

$$[VP'(x_0)](x) = VU(x),$$

*) La seule hypothèse d'existence de la dérivée est insuffisante.

quant au second, il tend vers zéro, puisque le rapport $\frac{\zeta(\Delta y)}{\|\Delta y\|}$ tend vers zéro et le facteur scalaire est borné. Donc

$$\lim_{t\to 0} \frac{R(x_0+tx)-R(x_0)}{t} = VU(x),$$

c.q.f.d.

Remarque. Si $P$ est dérivable en $x_0$, $R$ l'est également.

V. Si dans les conditions de IV, l'opérateur $Q = V$ est continu et linéaire, alors

$$R'(x_0) = VP'(x_0).$$

**1.3.** Signalons maintenant une inégalité qui dans le cas général remplace la formule des accroissements finis pour les fonctions réelles ordinaires.

Soient $x_0, x \in \Omega$; supposons que tous les points de l'intervalle $[x_0, x]$ appartiennent à $\Omega$ et qu'en chaque point de cet intervalle existe la dérivée de l'opérateur $P$. Posons

$$\varphi(t) = g(P(x_0 + t\,\Delta x)), \tag{4}$$

où $\Delta x = x - x_0$, $g$ est une fonctionnelle arbitraire de $\mathbf{Y}^*$.

Il est aisé de voir que la fonction réelle $\varphi$ possède une dérivée dans l'intervalle $[0, 1]$. En effet,

$$\frac{\varphi(t+\Delta t)-\varphi(t)}{\Delta t} = g\left(\frac{P(x_0+t\Delta x+\Delta t\,\Delta x)-P(x_0+t\,\Delta x)}{\Delta t}\right). \tag{5}$$

L'expression comprise sous le signe de la fonctionnelle $g$ tend vers $P'(x_0 + t\,\Delta x)(\Delta x)$ pour $\Delta t \to 0$, donc la fonctionnelle $g$ étant continue, le second membre de (5) a pour limite

$$\varphi'(t) = g(P'(x_0 + t\,\Delta x)(\Delta x)). \tag{6}$$

Ecrivons pour la fonction $\varphi$ la formule des accroissements finis

$$\varphi(1) - \varphi(0) = \varphi'(\theta) \quad (0 < \theta < 1).$$

Compte tenu de (4) et (6), on en déduit que

$$g(P(x_0 + \Delta x) - P(x_0)) = g(P'(x_0 + \theta\,\Delta x)(\Delta x)),$$

donc

$$|g(P(x_0+\Delta x)-P(x_0))| \leqslant \|g\| \sup_{0<\theta<1} \|P'(x_0+\theta\,\Delta x)\|\,\|\Delta x\|. \tag{7}$$

Prenons maintenant une fonctionnelle $g \neq 0$ telle que

$$g(P(x_0+\Delta x)-P(x_0)) = \|g\|\,\|P(x_0+\Delta x)-P(x_0)\|.$$

Il résulte alors de (7)

$$\|P(x)-P(x_0)\| \leqslant \sup_{0<\theta<1} \|P'(x_0+\theta\,\Delta x)\|\,\|\Delta x\| \tag{8}$$

$$(\Delta x = x - x_0).$$

En appliquant la majoration (8) à l'opérateur $P - U$ ($U \in B(\mathrm{X}, \mathrm{Y})$) et en utilisant les propositions II et III de 1.2, on est conduit à l'inégalité

$$\| P(x) - P(x_0) - U(\Delta x) \| \leqslant \sup_{0<\theta<1} \| P'(x_0 + \theta \Delta x) - U \| \, \| \Delta x \|.$$

En particulier, si l'on pose $U = P'(x_0)$, on obtient

$$\| P(x_0 + \Delta x) - P(x_0) - P'(x_0) \Delta x \| \leqslant \\ \leqslant \| \Delta x \| \sup_{0<\theta<1} \| P'(x_0 + \theta \Delta x) - P'(x_0) \|. \quad (9)$$

La relation (8) sera dite *formule des accroissements finis* et l'inégalité (9), *formule des accroissements finis à résidu.*

**1.4.** Voici quelques applications des formules des accroissements finis.

Supposons que l'opérateur $P$ admet une dérivée en chaque point d'un ensemble ouvert $\Omega_0 \subset \Omega$. Ceci définit un opérateur $P'$ associant à un élément $x \in \Omega_0$ l'élément $P'(x) \in B(\mathrm{X}, \mathrm{Y})$. Prouvons que si l'opérateur $P'$ est continu en un point $x_0 \in \Omega_0$, $P$ sera dérivable en ce point.

En effet, l'opérateur $P'$ étant par hypothèse continu, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $\| P'(x) - P'(x_0) \| < \varepsilon$ dès que $\| x - x_0 \| < \delta$, c'est-à-dire si $\| \Delta x \| < \delta$, alors $\sup\limits_{0<\theta<1} \| P'(x_0 + \theta \Delta x) - P'(x_0) \| \leqslant \varepsilon$. En appliquant ceci à (9) on obtient

$$\| P(x_0 + \Delta x) - P(x_0) - P'(x_0)(\Delta x) \| \leqslant \varepsilon \| \Delta x \|,$$

qui exprime que $P$ est dérivable au point $x_0$.

La formule des accroissements finis peut s'avérer utile pour démontrer que l'opérateur $P$ présente un point fixe.

THÉORÈME 1. *Supposons qu'un opérateur $P$ envoie l'ensemble $\Omega \subset \mathrm{X}$ dans $\mathrm{X}$ et possède une dérivée en tout point d'un ensemble fermé convexe $\Omega_0 \subset \Omega$. Si*

1) $P(\Omega_0) \subset \Omega$;
2) $\sup\limits_{x \in \Omega_0} \| P'(x) \| = \alpha < 1$,

*alors l'opérateur $P$ possède un point fixe unique contenu dans $\Omega_0$.*

DÉMONSTRATION. Il suffit de vérifier dans cette situation que les conditions du théorème XVI.1.1 sont satisfaites, c'est-à-dire $P$ est un opérateur de contraction.

Soient $x_1, x_2 \in \Omega_0$. La formule des accroissements finis donne

$$\| P(x_2) - P(x_1) \| \leqslant \| x_2 - x_1 \| \sup_{0<\theta<1} \| P'(x_1 + \\ + \theta (x_2 - x_1)) \| \leqslant \alpha \| x_2 - x_1 \|,$$

c.q.f.d.

Remarque. Si l'on suppose que l'opérateur $P'$ est continu dans $\Omega_0$, la condition 2) est une condition nécessaire pour que $P$ soit opérateur de contraction.

En effet, si

$$\sup_{x\in\Omega_0} \| P'(x) \| \geqslant 1,$$

il existe une suite numérique $\alpha_n \to 1$ et une suite $\{x_n\}$ d'éléments de $\Omega_0$, telles que

$$\| P'(x_n) \| > \alpha_n \quad (n = 1, 2, \ldots);$$

ceci étant on peut admettre que les points $x_n$ sont intérieurs à $\Omega_0$.

Par définition de la dérivée, pour tout $n = 1, 2, \ldots$ il existe un élément $x'_n$ tel que pour $t$ assez petits

$$\left\| \frac{P(x_n + tx'_n) - P(x_n)}{t} \right\| > \alpha_n \| x'_n \|$$

$$(|t| \leqslant t_n;\ n = 1, 2, \ldots).$$

En posant $y_n = x_n + t_n x'_n$, on obtient une suite de couples d'éléments $x_n$, $y_n$, telle que

$$\frac{\| P(y_n) - P(x_n) \|}{\| y_n - x_n \|} > \alpha_n, \tag{10}$$

c'est-à-dire $P$ n'est pas opérateur de contraction.

1.5. Voyons la signification des notions introduites plus haut lorsque l'un ou les deux espaces X et Y sont de dimension finie.

Supposons tout d'abord que X est un espace réel à une dimension, dont les éléments sont identifiés à des nombres.

Soit $U \in B(\mathrm{X}, \mathrm{Y})$. Il est immédiat de voir que dans ce cas $U$ s'écrit

$$U(t) = ty_0 \quad (t \in \mathrm{X}), \tag{11}$$

où $y_0$ est un élément de Y. De plus, puisque $\| U(t) \| = |t| \, \| y_0 \|$,

$$\| U \| = \| y_0 \|. \tag{12}$$

Au contraire, la formule (11) nous donne un opérateur $U \in B(\mathrm{X}, \mathrm{Y})$ quel que soit $y_0 \in \mathrm{Y}$. Il est clair que la correspondance entre les éléments des espaces Y et $B(\mathrm{X}, \mathrm{Y})$ est biunivoque, linéaire et, en vertu de (12), conserve la norme. Donc, on peut admettre que $B(\mathrm{X}, \mathrm{Y}) = \mathrm{Y}$.

Considérons maintenant un opérateur $F$ de $\Omega \subset \mathrm{X}$ dans Y. L'opérateur $F$ est donc une fonction de l'argument numérique à valeurs dans Y.

L'existence de la dérivée $F'(t_0) = y_0 \in \mathrm{Y}$ signifie que pour tout $t \in \mathrm{X}$

$$\lim_{\tau\to 0} \frac{F(t_0 + \tau t) - F(t_0)}{\tau} = ty_0. \tag{13}$$

En posant $\Delta t = \tau t$, on peut mettre (13) sous la forme

$$\lim_{\Delta t \to 0} \frac{F(t_0+\Delta t)-F(t_0)}{\Delta t} = y_0 .$$

Donc, dans le cas considéré, la définition de $F'(t_0)$ est celle d'une dérivée ordinaire d'une fonction réelle d'un argument réel. A noter que l'opérateur $F'$, comme $F$, envoie un ensemble de nombres dans **Y**, puisque $F'(t_0) \in \mathbf{Y}$.

Comme pour les fonctions réelles, on démontre sans peine que l'existence de la dérivée $F'(t_0)$ exprime la différentiabilité de l'opérateur $F$ en $t_0$.

Si $\Omega = [a, b]$, il semble naturel de considérer l'ensemble $F(\Omega)$ comme une *courbe* dans l'espace **Y**. L'élément $F'(t_0)$ donne alors la direction de la tangente à cette courbe en $y_0 = F(t_0)$.

**1.6.** Supposons maintenant que **X** et **Y** sont des espaces de dimension finie. Soient $m$ la dimension de **X**, $n$ celle de **Y**. Comme indiqué dans V.2.8, tome 1, tout opérateur $U \in B(\mathbf{X}, \mathbf{Y})$ est défini par une matrice rectangulaire

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \tag{14}$$

au moyen des formules

$$y = U(x) \quad (y = (\eta_1, \eta_2, \dots, \eta_n) \in \mathbf{Y},\ x = (\xi_1, \xi_2, \dots, \xi_m) \in \mathbf{X}),$$

$$\eta_i = \sum_{k=1}^{m} a_{ik}\xi_k \quad (i = 1, 2, \dots, n). \tag{15}$$

Soit un opérateur $P$ de $\Omega \subset \mathbf{X}$ dans **Y**. La donnée de l'opérateur $P$ équivaut à la donnée de $n$ fonctions numériques $\varphi_1, \varphi_2, \dots, \varphi_n$ de $m$ variables, de sorte que si

$$y = P(x) \quad (y = \eta_1, \eta_2, \dots, \eta_n) \in \mathbf{Y},\ x = (\xi_1, \xi_2, \dots, \xi_m) \in \Omega,$$

alors

$$\eta_i = \varphi_i(\xi_1, \xi_2, \dots, \xi_m) \quad (i = 1, 2, \dots, n). \tag{16}$$

Supposons qu'existe la dérivée $P'(x_0) = U$ $(x_0 = (\xi_1^{(0)}, \xi_2^{(0)}, \dots \dots, \xi_m^{(0)}))$ et que $U$ est défini par la matrice (14). En détaillant

$$\lim_{t \to 0} \frac{P(x_0+tx)-P(x_0)}{t} = U(x)$$

et en tenant compte de (15) et (16) on obtient $n$ relations

$$\lim_{t \to 0} \frac{\varphi_i(\xi_1^{(0)}+t\xi_1, \dots, \xi_m^{(0)}+t\xi_m)-\varphi_i(\xi_1^{(0)}, \dots, \xi_m^{(0)})}{t} = \sum_{k=1}^{m} a_{ik}\xi_k$$

$$(i = 1, 2, \dots, n). \tag{17}$$

Ces égalités étant satisfaites pour tous les $x \in X$, en prenant dans chacune d'elles successivement des $x$ dont toutes les coordonnées sont nulles à l'exception d'une qui est égale à l'unité, on trouve que les fonctions $\varphi_1, \varphi_2, \ldots, \varphi_n$ possèdent des dérivées partielles par rapport à $\xi_1, \xi_2, \ldots, \xi_m$, soit

$$\frac{\partial \varphi_i (\xi_1^{(0)}, \ldots, \xi_m^{(0)})}{\partial \xi_k} = a_{ik} \quad (i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m).$$

Donc, la dérivée $P'(x_0)$ est un opérateur linéaire, défini par la matrice des dérivées partielles des fonctions $\varphi_1, \varphi_2, \ldots, \varphi_n$.

A noter toutefois que l'existence des dérivées partielles des fonctions $\varphi_1, \varphi_2, \ldots, \varphi_n$ n'assure pas l'existence de la dérivée $P'(x_0)$ comme le prouve l'exemple suivant: soient $n = 1$, $m = 2$,

$$\varphi_1 (\xi_1, \xi_2) = \frac{\xi_1 \xi_2}{(\xi_1^2 + \xi_2^2)^2}, \quad \varphi_1 (0, 0) = 0;\ x_0 = (0, 0).$$

Il est clair que $\dfrac{\partial \varphi_1 (0, 0)}{\partial \xi_1} = \dfrac{\partial \varphi_1 (0, 0)}{\partial \xi_2} = 0$. Donc, si la dérivée $P'(x_0)$ existait, elle serait l'opérateur nul et la relation (17) aurait entraîné

$$\lim_{t \to 0} \frac{\varphi_1 (t\xi_1, t\xi_2)}{t} = 0.$$

Or, cette limite est en réalité égale à l'infini, si seulement $\xi_1 \neq 0$ et $\xi_2 \neq 0$.

On laisse au lecteur le soin de prouver qu'une condition nécessaire et suffisante pour que l'opérateur $P$ soit différentiable est que les fonctions $\varphi_1, \varphi_2, \ldots, \varphi_n$ admettent des différentielles. Signalons que dans le cas général la différentiabilité de l'opérateur $P$ ne résulte pas de l'existence de la dérivée comme c'est le cas en analyse élémentaire.

1.7. Considérons une fonction $F$ à variable réelle prenant ses valeurs dans un $B$-espace X. Si $F$ est définie sur l'intervalle $[a, b]$, on peut envisager son intégrale comme étant la limite des sommes intégrales

$$\sum_{k=0}^{n-1} F(\tau_k)(t_{k+1} - t_k)$$

$$(a = t_0 < t_1 < \ldots < t_n = b;\ \tau_k \in [t_k, t_{k+1}];\ k = 0, 1, \ldots, n-1)$$

pour $\lambda = \max_k [t_{k+1} - t_k] \to 0$. Si cette limite existe, on l'appelle *intégrale* de la fonction $F$ et on la note $\int_a^b F(t)\, dt$. Comme pour les fonctions à valeurs réelles, on démontre qu'une fonction continue est uniformément continue, donc intégrable, c'est-à-dire son intégrale existe. Les propriétés classiques de l'intégrale de Riemann s'étendent avec leurs démonstrations à l'intégrale abstraite définie plus haut. Signalons trois d'entre elles.

I. Si $U \in B(X, Y)$, alors

$$\int_a^b U(F(t))\,dt = U\left(\int_a^b F(t)\,dt\right).$$

II. Si $F(t) = \varphi(t)\, x_0$ ($t \in [a, b]$), où $x_0 \in X$ est un élément fixe et $\varphi(t)$ une fonction réelle intégrable, alors

$$\int_a^b F(t)\,dt = x_0 \int_a^b \varphi(t)\,dt.$$

III. $$\left\| \int_a^b F(t)\,dt \right\| \leqslant \int_a^b \| F(t) \|\,dt.$$

Revenons maintenant aux opérateurs non linéaires.

Soit un opérateur $R$ défini sur l'intervalle $[x_0, x_0 + \Delta x]$ ($x_0, x_0 + \Delta x \in X$) à valeurs dans l'espace $B(X, Y)$. On convient que

$$\int_{x_0}^{x_0+\Delta x} R(x)\,dx = \int_0^1 R(x_0 + t\Delta x)\,\Delta x\,dt =$$
$$= \lim \sum_{k=0}^{n-1} R(x_0 + \tau_k \Delta x)\,\Delta x\,(t_{k+1} - t_k).$$

Il est évident que si $R$ est continu, l'intégrale existe et est élément de Y. Les intégrales de ce type ont été introduites par M. Gavourine [1].

En particulier, si $R = P'$, où $P$ est un opérateur de $\Omega \subset X$ dans Y, possédant une dérivée $P'$ continue sur l'intervalle $[x_0, x_0 + \Delta x] \subset \Omega$, alors l'intégrale de $P'(x)$ existe et nous allons montrer qu'est vraie la relation

$$\int_{x_0}^{x_0+\Delta x} P'(x)\,dx = P(x_0 + \Delta x) - P(x_0), \tag{18}$$

qui généralise le théorème fondamental du calcul différentiel: la formule de Newton-Leibniz.

En effet,

$$\int_{x_0}^{x_0+\Delta x} P'(x)\,dx = \lim_{\lambda \to 0} \sum_{h=0}^{n-1} P'(x_0 + \tau_k \Delta x)\,\Delta x\,(t_{k+1} - t_k) =$$
$$= \lim_{\lambda \to 0} \sum_{k=0}^{n-1} P'(\overline{x}_k)\,\Delta x_k$$

$$(\overline{x}_k = x_0 + \tau_k \Delta x;\ \ \Delta x_k = (t_{k+1} - t_k)\,\Delta x\,;\ k = 0, \ldots, n-1).$$

Par ailleurs

$$P(x_0+\Delta x)-P(x_0)=\sum_{k=0}^{n-1}[P(x_0+t_{k+1}\Delta x)-P(x_0+t_k\Delta x)]=$$

$$=\sum_{k=0}^{n-1}[P(x_{k+1})-P(x_k)].$$

La formule des accroissements finis donne

$$\left\|\sum_{k=0}^{n-1}[P(x_{k+1})-P(x_k)-P'(\bar{x}_k)\Delta x_k]\right\|\leqslant$$

$$\leqslant\|\Delta x\|\sum_{k=0}^{n-1}(t_{k+1}-t_k)\sup_{0<\theta<1}\|P'(x_k+\theta\,\Delta x_k)-P'(\bar{x}_k)\|,$$

d'où, compte tenu de la continuité de $P'$, donc de sa continuité uniforme sur l'intervalle $[x_0,\ x_0+\Delta x]$, l'on déduit le résultat annoncé.

REMARQUE. De la propriété III de l'intégrale il résulte que si

$$\|R(x_0+\tau\,\Delta x)\|\leqslant\varphi(t_0+\tau\,\Delta t)\quad(\tau\in[0,\ 1])\tag{19}$$

et

$$\|\Delta x\|\leqslant\Delta t,\tag{20}$$

alors

$$\left\|\int_{x_0}^{x_0+\Delta x}R(x)\,dx\right\|\leqslant\int_{t_0}^{t_0+\Delta t}\varphi(t)\,dt.\tag{21}$$

En effet,

$$\left\|\int_{x_0}^{x_0+\Delta x}R(x)\,dx\right\|=\left\|\int_0^1 R(x_0+\tau\Delta x)(\Delta x)\,d\tau\right\|\leqslant$$

$$\leqslant\|\Delta x\|\int_0^1\|R(x_0+\tau\Delta x)\|\,d\tau\leqslant\Delta t\int_0^1\varphi(t_0+\tau\Delta t)\,d\tau=$$

$$=\int_{t_0}^{t_0+\Delta t}\varphi(t)\,dt.$$

De (21) il suit en particulier que, si l'inégalité

$$\|R(x)\|\leqslant\varphi(t)\tag{22}$$

est réalisée pour tous les $x$ et $t$ tels que

$$\|x-x_0\|\leqslant t-t_0,\tag{23}$$

alors

$$\left\|\int_{x_0}^{x_1}R(x)\,dx\right\|\leqslant\int_{t_0}^{t_1}\varphi(t)\,dt,\tag{24}$$

où $x_1$ est un élément quelconque vérifiant la condition $\| x_1 - x_0 \| \leqslant$ $\leqslant t_1 - t_0$.

En effet, en posant

$$\Delta x = x_1 - x_0, \quad \Delta t = t_1 - t_0, \quad x = x_0 + \tau \, \Delta x, \quad t = t_0 + \tau \, \Delta t$$

on trouve que $\| \Delta x \| \leqslant \Delta t$ et

$$\| x - x_0 \| = \tau \, \| \Delta x \| \leqslant \tau \, \Delta t = t - t_0 \quad (\tau \in [0, 1]),$$

c'est-à-dire est réalisée (23) et, par suite de (22),

$$\| R (x_0 + \tau \, \Delta x) \| \leqslant \varphi (t_0 + \tau \, \Delta t) \quad (\tau \in [0, 1]).$$

En se servant de (21), on trouve (24).

### § 2. Dérivée seconde et opérateurs bilinéaires

**2.1.** Supposons que dans un ensemble ouvert $\Omega$ d'un $B$-espace **X** existe la dérivée $P'$ d'un opérateur $P$ de $\Omega$ dans un $B$-espace **Y**. On a indiqué plus haut que $P'$ peut être traitée comme un opérateur de $\Omega$ dans l'espace $B$ (**X**, **Y**). On peut donc envisager la dérivée (si elle existe) de $P'$ en un point $x_0$ de $\Omega$, que l'on appelle *dérivée seconde* de $P$ et l'on note $P''(x_0)$. Si l'opérateur $P'$ est dérivable, on dit que $P$ est *deux fois dérivable*.

Si la dérivée seconde existe en chaque point de l'ensemble $\Omega$, elle définit l'opérateur $P''$ dont la dérivée est appelée *dérivée troisième* de $P$. D'une façon générale, la *dérivée* $P^{(n)}(x_0)$ *d'ordre n* en $x_0$ est, par définition, la dérivée de l'opérateur $P^{(n-1)}$. Il est clair que

$$P^{(n)}(x_0) \in B \overbrace{(\mathbf{X}, B(\mathbf{X}, \ldots, B(\mathbf{X}}^{n}, \mathbf{Y}), \ldots)).$$

**2.2.** L'interprétation directe des éléments de l'espace $B$ (**X**, $B$ (**X**, **Y**)) est assez confuse. Introduisons une nouvelle notion pour les définir de façon suggestive.

Soient donnés deux $B$-espaces **X** et **Y** et supposons qu'à tout couple ordonné d'éléments $x, x' \in \mathbf{X}$ est associé l'élément

$$y = B(x, x') \in \mathbf{Y}.$$

L'opérateur binaire $B$ est *bilinéaire* si sont réunies les deux conditions suivantes.

I. Quels que soient $x_1, x_2, x_1', x_2' \in \mathbf{X}$ et les nombres $\alpha, \beta$, on a

$$\left. \begin{aligned} B(\alpha x_1 + \beta x_2, x') &= \alpha B(x_1, x') + \beta B(x_2, x'), \\ B(x, \alpha x_1' + \beta x_2') &= \alpha B(x, x_1') + \beta B(x, x_2') \end{aligned} \right\} \quad (x, x' \in \mathbf{X}). \qquad (1)$$

II. Il existe un réel $M > 0$ tel que pour tous $x, x' \in \mathbf{X}$

$$\| B(x, x') \| \leqslant M \, \| x \| \, \| x' \|. \qquad (2)$$

Comme pour les opérateurs linéaires, la plus petite valeur de $M$ dans l'inégalité (2) est appelée *norme* de l'opérateur $B$ et notée comme toujours $\|B\|$. Il est clair que

$$\|B\| = \sup_{\|x\|,\ \|x'\| \leqslant 1} \|B(x, x')\|. \tag{3}$$

Il est immédiat de vérifier que l'ensemble des opérateurs bilinéaires de norme (3), linéarisé de façon naturelle, est un $B$-espace que l'on notera $B(\mathbf{X}^2, \mathbf{Y})$.

Sans nous attarder sur les détails, indiquons que l'on introduit de façon analogue les espaces $B(\mathbf{X}^n, \mathbf{Y})$ des opérateurs $n$-linéaires *).

Considérons quelques exemples d'opérateurs bilinéaires.

Dans le cas élémentaire où $\mathbf{X} = \mathbf{Y} = \mathbf{R}^1$, la forme générale des opérateurs bilinéaires est

$$B(x, x') = \alpha x x'$$

($\alpha$ est un réel).

Supposons maintenant que $\mathbf{X}$ et $\mathbf{Y}$ sont des espaces de dimension finie, $\mathbf{X}$ de dimension $m$, $\mathbf{Y}$ de dimension $n$. Appelons $x_i$ ($i = 1, 2, \ldots, m$) l'élément de $\mathbf{X}$ dont toutes les coordonnées sont nulles hormis la coordonnée $i$-ième qui est égale à l'unité. Il en va de même pour $y_j \in \mathbf{Y}$ ($j = 1, 2, \ldots, n$). Soit un opérateur $B \in B(\mathbf{X}^2, \mathbf{Y})$. En posant

$$B(x_i, x_j) = (a_{ij}^{(1)}, a_{ij}^{(2)}, \ldots, a_{jj}^{(n)}) \quad (i, j = 1, 2, \ldots, m),$$

en vertu de (1) on obtient pour des éléments arbitraires $x$ et $x'$ de $\mathbf{X}$ $(x = (\xi_1, \xi_2, \ldots, \xi_m),\quad x' = (\xi_1', \xi_2', \ldots, \xi_m'))$

$$y = B(x, x') = B\Big(\sum_{i=1}^{m} \xi_i x_i,\ \sum_{j=1}^{m} \xi_j' x_j\Big) = \sum_{i,j=1}^{m} \xi_i \xi_j' B(x_i, x_j).$$

Si donc $y = (\eta_1, \eta_2, \ldots, \eta_n)$, alors

$$\eta_k = \sum_{i,j=1}^{m} a_{ij}^{(k)} \xi_i \xi_j' \quad (k = 1, 2, \ldots, n), \tag{4}$$

c'est-à-dire les coordonnées de l'élément $y$ sont des formes bilinéaires des coordonnées des éléments $x$ et $x'$.

Il est clair qu'un opérateur $B$ défini par les formules (4) sera bilinéaire quelle que soit la matrice à trois entrées

$$(a_{ij}^{(k)}) \quad (i, j = 1, 2, \ldots, m;\quad k = 1, 2, \ldots, n). \tag{5}$$

Donc, tout opérateur bilinéaire définit et est lui-même défini par une matrice à trois entrées (5).

---

*) Un exposé détaillé de la théorie des opérateurs $n$-linéaires est accessible dans Gavourine [2], Hille et Phillips.

S'agissant de la norme de l'opérateur $B$, elle dépend de toute évidence des normes des espaces X et Y. Si $X = l^2_m$, $Y = l^2_n$, il s'ensuit de (4) grâce à l'inégalité de Cauchy-Bouniakovski

$$|\eta_k| = \Big|\sum_{i,j=1}^{m} a_{ij}^{(k)}\xi_i\xi'_j\Big| \leqslant$$
$$\leqslant \Big[\sum_{i=1}^{m} |\xi|^2\Big]^{1/2} \Big[\sum_{i=1}^{m} \Big|\sum_{j=1}^{m} a_{ij}^{(k)}\xi'_j\Big|^2\Big]^{1/2} \leqslant \Lambda_k \, \|x\| \, \|x'\|,$$

où $\Lambda_k^2$ désigne la plus grande valeur propre de la matrice $A_k A_k^*$, produit de la matrice

$$A_k = \begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \ldots & a_{1m}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} & \ldots & a_{2m}^{(k)} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1}^{(k)} & a_{m2}^{(k)} & \ldots & a_{mm}^{(k)} \end{pmatrix} \quad (k = 1, 2, \ldots, n)$$

par sa transposée $A_k^*$ (cf. V.2.8, tome 1). Donc

$$\|y\| = \Big[\sum_{k=1}^{n} |\eta_k|^2\Big]^{1/2} \leqslant \Big[\sum_{k=1}^{n} \Lambda_k^2\Big]^{1/2} \|x\| \, \|x'\|.$$

D'où

$$\|B\| \leqslant \Big[\sum_{k=1}^{n} \Lambda_k^2\Big]^{1/2}. \tag{6}$$

Comme

$$\Lambda_k^2 \leqslant \sum_{i,j=1}^{m} |a_{ij}^{(k)}|^2 \quad (k = 1, 2, \ldots, n),$$

de (6) on déduit une majoration moins bonne mais plus simple

$$\|B\| \leqslant \Big[\sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} |a_{ij}^{(k)}|^2\Big]^{1/2}.$$

Si $X = l^\infty_m$, $Y = l^\infty_n$, des raisonnements évidents nous conduisent à une majoration analogue pour $\|B\|$:

$$\|B\| \leqslant \max_k \sum_{i=1}^{m} \sum_{j=1}^{m} |a_{ij}^{(k)}|.$$

Les opérateurs bilinéaires présentent un plus grand intérêt dans les espaces de dimension infinie et particulièrement dans les espaces de fonctions.

L'opérateur $B$:

$$y = B(x, x'), \quad y(s) = \int_0^1 \int_0^1 K(s, t, u)\, x(t)\, x'(u)\, dt\, du \tag{7}$$

sera bilinéaire dans tel ou tel espace de fonctions (définies sur [0, 1]) selon les conditions imposées à son noyau $K$. Si l'on suppose que la fonction $K$ est continue, $B$ sera bilinéaire quels que soient les espaces $\mathbf{X}$ et $\mathbf{Y}$ pris parmi les espaces $\mathbf{C}$, $\mathbf{L}^p$ $(1 \leqslant p \leqslant \infty)$.

Cependant, les opérateurs de la forme

$$y = B(x, x') \tag{8}$$

$$y(s) = \int_0^1 K(s, t)\, x(t)\, x'(t)\, dt \tag{9}$$

sont les plus intéressants dans les applications. Supposons tout d'abord que $\mathbf{X} = \mathbf{Y} = \mathbf{L}^p(0, 1)$ $(p \geqslant 2)$. Si $M < \infty$, où

$$M^p = \begin{cases} \int_0^1 \left[ \int_0^1 |K(s, t)|^{\frac{p}{p-2}}\, dt \right]^{p-2} ds & (p > 2), \\ \int_0^1 \sup_s |K(s, t)|\, dt & (p = 2), \end{cases}$$

alors $B \in B(\mathbf{X}^2, \mathbf{Y})$.

En effet, le produit figurant sous le signe somme dans (9) est sommable pour presque tous les $s \in [0, 1]$. Pour s'en assurer dans le cas où $p > 2$, il suffit d'appliquer à l'intégrale (9) l'inégalité généralisée de Hölder (II.3.4, tome 1) avec les puissances $\frac{p}{p-2}$, $p$, $p$. Cette inégalité nous dit que

$$|y(s)|^p = \left| \int_0^1 K(s, t)\, x(t)\, x'(t)\, dt \right|^p \leqslant$$

$$\leqslant \left[ \int_0^1 |K(s, t)|^{\frac{p}{p-2}}\, dt \right]^{p-2} \int_0^1 |x(t)|^p\, dt \int_0^1 |x'(t)|^p\, dt$$

et par suite

$$\int_0^1 |y(s)|^p\, ds \leqslant M^p \|x\|^p \|x'\|^p < \infty. \tag{10}$$

Donc $y \in \mathbf{L}^p(0, 1)$. On a démontré par là même que l'inégalité (2) est réalisée pour l'opérateur $B$. L'opérateur $B$ est bilinéaire, car

additif (relation (1)). De (10) on déduit que

$$\| B \| \leqslant M.$$

Nous avons traité le cas $p > 2$. Le cas $p = 2$ ne présente pas de difficultés.

Supposons maintenant que X = Y = C [0, 1]. Pour que l'opérateur (8) soit bilinéaire, il faut que la fonction $K(s, t)$ soit continue en $s$ et que

$$M = \max_s \int_0^1 |K(s, t)|\, dt < \infty.$$

En effet, la sommabilité de la fonction à intégrer dans (9) ne fait aucun doute. La continuité de $y$ découle de la possibilité de passer à la limite sous le signe somme dans (9). Les égalités (1) sont évidentes. On vérifie sans peine que l'inégalité (2) est réalisée, puisque

$$|y(s)| \leqslant \max_t |x(t)| \max_t |x'(t)| \int_0^1 |K(s, t)|\, dt \quad (0 \leqslant s \leqslant 1),$$

et par suite

$$\| y \| \leqslant M \| x \| \| x' \|.$$

D'où

$$\| B \| \leqslant M = \sup_s \int_0^1 |K(s, t)|\, dt.$$

En fait on a le signe d'égalité

$$\| B \| = \sup_s \int_0^1 |K(s, t)|\, dt$$

(on établit ceci exactement comme l'égalité respective pour les opérateurs linéaires, V.2.4, tome 1).

2.3. L'espace $B(X^2, Y)$ est intimement lié à l'espace $B(X, B(X, Y))$ que nous avons rencontré en étudiant la dérivée seconde. Plus exactement, nous allons montrer que ces deux espaces sont linéairement isométriques.

Soit $W \in B(X, B(X, Y))$. Prenons un élément quelconque $x' \in X$ et posons $U_{x'} = W(x')$. De toute évidence, $U_{x'} \in B(X, Y)$, si bien que $y = B(x, x') = U_{x'}(x)$ est élément de Y. L'opérateur binaire $B$ ainsi défini vérifie la condition I de 2.2 manifestement, et la condition II, puisque

$$\| y \| \leqslant \| U_{x'} \| \| x \| \leqslant \| W \| \| x \| \| x' \|. \tag{11}$$

Ainsi, $B \in B(\mathbf{X}^2, \mathbf{Y})$ et de plus (11) entraîne

$$\|B\| \leqslant \|W\|. \tag{12}$$

Donc, à tout opérateur $W \in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ on peut associer par ce procédé un opérateur bilinéaire $B \in B(\mathbf{X}^2, \mathbf{Y})$. Montrons que chaque opérateur bilinéaire $B \in B(\mathbf{X}^2, \mathbf{Y})$ est l'image par cette correspondance d'un opérateur $W \in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$.

Il suffit de remarquer à cet effet que si $B$ est donné, alors pour $x' \in \mathbf{X}$ fixe l'opérateur *)

$$W(x') = B(\cdot, x') \in B(\mathbf{X}, \mathbf{Y}).$$

De plus

$$\|W(x')\| = \sup_{\|x\| \leqslant 1} \|W(x')(x)\| = \sup_{\|x\| \leqslant 1} \|B(x, x')\| \leqslant \|B\| \; \|x'\|,$$

d'où il suit que $W \in B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ et que

$$\|W\| \leqslant \|B\|.$$

En combinant cette inégalité avec (12), on obtient l'égalité

$$\|W\| = \|B\|. \tag{13}$$

La correspondance entre les éléments des espaces $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ et $B(\mathbf{X}^2, \mathbf{Y})$ étant additive et homogène, il s'ensuit de (13) qu'elle est biunivoque, donc est une isométrie linéaire.

L'isométrie linéaire des espaces $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$ et $B(\mathbf{X}^2, \mathbf{Y})$ nous permet d'identifier les éléments associés de ces espaces, ce que du reste nous ferons dans la suite.

Signalons sans le démontrer (la démonstration facile est laissée au soin du lecteur) que dans le cas général l'espace $B(\mathbf{X}, B(\mathbf{X}, \ldots$ $\ldots, B(\mathbf{X}, \mathbf{Y}) \ldots))$ et l'espace $B(\mathbf{X}^n, \mathbf{Y})$ des opérateurs $n$-linéaires sont linéairement isométriques (cf. Gavourine [2]).

2.4. Comme indiqué dans 2.1, la dérivée seconde de l'opérateur $P$ de $\mathbf{X}$ dans $\mathbf{Y}$ est un élément de l'espace $B(\mathbf{X}, B(\mathbf{X}, \mathbf{Y}))$, donc, d'après les résultats précédents, $P''(x_0)$ peut être traitée comme un opérateur bilinéaire. Ceci étant, il vient en vertu de 2.3

$$\lim_{t \to 0} \frac{P'(x_0 + tx') - P'(x_0)}{t} = P''(x_0)(\cdot, x'). \tag{14}$$

Donc, pour tout $x \in \mathbf{X}$

$$P''(x_0)(x, x') = \lim_{t \to 0} \frac{P'(x_0 + tx')\,x - P'(x_0)\,x}{t}. \tag{15}$$

A noter que les égalités (14) et (15) ne sont pas équivalentes. De façon plus précise, il peut exister un opérateur bilinéaire $B \in$

*) $B(\cdot, x')$ désigne un opérateur $U$ tel que $U(x) = B(x, x')$. On se servira de notations analogues dans la suite.

$\in B\,(X^2,\ Y)$ tel que pour tous $x,\ x' \in X$

$$\lim_{t\to 0} \frac{P'(x_0+tx')\,x - P'(x_0)\,x}{t} = B\,(x,\ x'), \tag{16}$$

alors que $P''\,(x_0)$ n'existe pas. Il est immédiat de voir qu'une condition nécessaire et suffisante d'existence de $P''\,(x_0)$ est que pour tout $x' \in X$ la limite dans (16) soit uniforme en $x \in X$, $\|x\| = 1$. Le fait signalé permet d'élargir un peu la notion de dérivée seconde. Plus exactement, on dira qu'un opérateur bilinéaire $B$ est *dérivée seconde faible* d'un opérateur $P$ en un point $x_0$ si (16) a lieu pour tous $x,\ x' \in X$. On peut de toute évidence conserver l'ancienne notation pour la dérivée faible, car si la dérivée seconde ordinaire existe, elle est confondue avec la faible.

REMARQUE. Il est aisé de voir que si $P$ est deux fois dérivable en $x_0$, c'est que (16) a lieu uniformément en $x,\ x' \in X$, $\|x\| = 1$, $\|x'\| = 1$.

Etablissons la forme de la dérivée seconde dans le cas où X et Y sont des espaces de dimension finie.

Supposons que $P''\,(x_0)$ existe et en tant qu'opérateur bilinéaire est définie par la matrice

$$(a_{ij}^{(k)}) \quad (i,\ j = 1,\ 2,\ \ldots,\ m;\ k = 1,\ 2,\ \ldots,\ n).$$

En identifiant les coordonnées respectives des deux membres de (15), on obtient dans les notations de 1.6

$$\sum_{i,j=1}^{m} a_{ij}^{(k)} \xi_i \xi_j' = \lim_{t\to 0} \frac{\sum\limits_{i=1}^{m} \frac{\partial \varphi_k}{\partial \xi_i}(\xi_j^{(0)} + t\xi_j')\,\xi_i - \sum\limits_{i=1}^{m} \frac{\partial \varphi_k}{\partial \xi_i}(\xi_j^{(0)})\,\xi_i}{t}$$

$$(k = 1,\ 2,\ \ldots,\ n)^{*)}.$$

Si pour $x$ et $x'$ on prend des éléments dont toutes les coordonnées sont nulles à l'exception respectivement de la $i$-ième et la $j$-ième qui sont égales à l'unité, on obtient que les fonctions $\varphi_k$ possèdent des dérivées secondes et

$$a_{ij}^{(k)} = \frac{\partial^2 \varphi_k\,(\xi_1^{(0)},\ \xi_2^{(0)},\ \ldots,\ \xi_m^{(0)})}{\partial \xi_i\,\partial \xi_j} \quad (i,\ j = 1,\ 2,\ \ldots,\ m;\quad k = 1,\ 2,\ \ldots,\ n).$$

2.5. On achève ce paragraphe par l'établissement d'une formule qui généralise celle de Taylor pour les fonctions réelles.

On suppose toujours que l'opérateur $P$ est un opérateur de $\Omega \subset X$ dans Y et qu'il admet une dérivée seconde continue sur l'intervalle

---

*) On s'est servi de la notation $\varphi\,(\xi_j) = \varphi\,(\xi_1,\ \xi_2,\ \ldots,\ \xi_m)$.

$[x_0, \bar{x}] \subset \Omega$. On a alors

$$P(\bar{x}) - P(x_0) - P'(x_0)(\bar{x} - x_0) = \int_{x_0}^{\bar{x}} P''(x)(\bar{x} - x, \cdot)\, dx. \quad (17)$$

Pour prouver ceci, on considère l'opérateur $Q$:

$$Q(x) = P(x) + P'(x)(\bar{x} - x) = P(x) + Q_0(x) \quad (x \in \Omega)$$

et on vérifie que pour tout $\tilde{x} \in [x_0, \bar{x}]$

$$Q'(\tilde{x})(x) = P''(\tilde{x})(\bar{x} - \tilde{x}, x) \quad (x \in X). \quad (18)$$

Ceci étant, la formule (17) découlera de toute évidence de la formule de Newton-Leibniz (cf. (18), 1.7).

Il suffit de considérer l'opérateur $Q_0$. On a

$$\frac{Q_0(\tilde{x} + tx) - Q_0(\tilde{x})}{t} = \frac{P'(\tilde{x} + tx)(\bar{x} - \tilde{x} - tx) - P'(\tilde{x})(\bar{x} - \tilde{x})}{t} =$$

$$= \frac{P'(\tilde{x} + tx)(\bar{x} - \tilde{x}) - P'(\tilde{x})(\bar{x} - \tilde{x})}{t} - P'(\tilde{x} + tx)\, x.$$

Pour $t \to 0$ le premier terme tend vers $P''(\tilde{x})(\bar{x} - \tilde{x}, x)$, le second, vers $P'(\tilde{x})(x)$, puisque $P'$ est manifestement continue. Donc,

$$Q_0'(\tilde{x})(x) = P''(\tilde{x})(\bar{x} - \tilde{x}, x) - P'(\tilde{x})\, x,$$

d'où l'on déduit (18).

REMARQUE. Il est clair que la formule (17) subsiste si par $P''(x)$ on comprend la dérivée seconde faible.

## § 3. Exemples

Dans les paragraphes précédents, nous avons eu affaire à des exemples de calcul des dérivées d'opérateurs non linéaires dans le cas le plus simple d'une dimension finie. Dans ce paragraphe, nous considérons des exemples plus compliqués et plus consistants, dont nous nous servirons dans le paragraphe suivant et essentiellement dans le chapitre suivant.

**3.1.** Soit l'opérateur intégral $P$:

$$y = P(x), \quad y(s) = \int_0^1 K(s, t, x(t))\, dt. \quad (1)$$

THÉORÈME 1. *Supposons que la fonction $K(s, t, u)$ est continue et bicontinûment dérivable par rapport à $s$, $t$ et $u$ pour $0 \leqslant s, t \leqslant 1$, $-\infty < u < \infty$, et que pour ces valeurs de $s$, $t$, $u$*

$$|K''_{u^2}(s, t, u)| \leqslant M\, |u|^{p-2} + N. \quad (2)$$

*Si $p \geqslant 2$, alors (1) est un opérateur différentiable de $L^p(0, 1)$ dans $L^q(0, 1)$ $(1 \leqslant q < \infty)$ admettant une dérivée seconde en chaque point $x_0 \in L^p(0, 1)$. Ceci étant, $P'(x_0) = U$, $P''(x_0) = B$, où $U$ et $B$ sont définis par les formules*

$$z = U(x), \qquad z(s) = \int_0^1 K'_u(s, t, x_0(t))\, x(t)\, dt,$$

$$v = B(x, x'), \quad v(s) = \int_0^1 K''_{u^2}(s, t, x_0(t))\, x(t)\, x'(t)\, dt.$$

DÉMONSTRATION. Assurons-nous tout d'abord que (1) est un opérateur de $L^p(0, 1)$ dans $L^\infty(0, 1)$ (donc dans $L^q(0, 1)$ $(q < \infty)$). On observera à cet effet que la fonction composée $K(s, t, x(t))$ est mesurable, puisque $K(s, t, u)$ est continue.

La formule de Taylor nous donne

$$K(s, t, x(t)) = K(s, t, 0) + K'_u(s, t, 0)\, x(t) + \\ + \frac{1}{2} K''_{u^2}(s, t, \theta x(t))\, x^2(t) \qquad (0 < \theta < 1).$$

Evaluons l'intégrale de chaque terme du second membre

$$\left| \int_0^1 K(s, t, 0)\, dt \right| \leqslant \max_{s, t} | K(s, t, 0) | = A_1,$$

$$\left| \int_0^1 K'_u(s, t, 0)\, x(t)\, dt \right| \leqslant \max_{s, t} | K'_u(s, t, 0) | \int_0^1 | x(t) |\, dt \leqslant A_2 \| x \|_{L^p}.$$

Servons-nous de (2) pour majorer le troisième terme:

$$\left| \frac{1}{2} \int_0^1 K''_{u^2}(s, t, \theta x(t))\, x^2(t)\, dt \right| \leqslant \\ \leqslant \frac{1}{2} \int_0^1 [M | x(t) |^{p-2} + N] | x(t) |^2\, dt \leqslant A_3 \| x \|^p_{L^p}.$$

Il s'ensuit que $y = P(x)$ a un sens et en outre

$$| y(s) | \leqslant A_1 + A_2 \| x \|_{L^p} + A_3 \| x \|^p_{L^p} \quad (s \in [0, 1]),$$

de sorte que $y \in L^\infty(0, 1)$.

Prouvons maintenant que $P$ est dérivable et que $P'(x_0) = U$. Supposons que $z = U(x)$ et que

$$z_\tau = \frac{P(x_0 + \tau x) - P(x_0)}{\tau}.$$

On a

$$| z_\tau(s) - z(s) | =$$

$$= \left| \int_0^1 \left[ \frac{K(s, t, x_0(t) + \tau x(t)) - K(s, t, x_0(t))}{\tau} - K'_u(s, t, x_0(t))\, x(t) \right] dt \right|. \tag{3}$$

Comme, d'après la formule de Taylor,

$$K(s, t, x_0(t) + \tau x(t)) = K(s, t, x_0(t)) +$$
$$+ \tau K'_u(s, t, x_0(t))\, x(t) + \frac{\tau^2}{2} K''_{u^2}(s, t, x_0(t) + \theta\tau x(t))\, x^2(t)$$
$$(0 < \theta < 1),$$

on peut mettre (3) sous la forme

$$| z_\tau(s) - z(s) | = \frac{|\tau|}{2} \left| \int_0^1 K''_{u^2}(s, t, x_0(t) + \theta\tau x(t))\, dt \right|.$$

En reprenant les raisonnements ci-dessus, on obtient la majoration

$$| z_\tau(s) - z(s) | \leqslant A_4 \, | \tau | \quad (0 \leqslant s \leqslant 1),$$

où la constante $A_4$ dépend non pas de $x$, mais de sa norme $\| x \|$. Donc, $z_\tau \to z$ pour $\tau \to 0$ uniformément en $x$, $\| x \| = 1$.

Passons à la dérivée seconde. Posons

$$v = B(x, x'); \quad v_\tau = \frac{P'(x_0 + \tau x')(x) - P'(x_0)(x)}{\tau}.$$

Comme indiqué dans 2.4, il suffit de prouver que pour $\tau \to 0$

$$v_\tau \to v$$

uniformément en $x \in L^p(0, 1)$, $\| x \| = 1$.

En appliquant la formule des accroissements finis et ensuite l'inégalité de Hölder, on obtient

$$| v_\tau(s) - v(s) | = \left| \int_0^1 \left[ \frac{K'_u(s, t, x_0(t) + \tau x'(t)) - K'_u(s, t, x_0(t))}{\tau} - \right.\right.$$
$$\left.\left. - K''_{u^2}(s, t, x_0(t))\, x'(t) \right] x(t)\, dt \right| =$$
$$= \left| \int_0^1 [K''_{u^2}(s, t, x_0(t) + \tau\theta x'(t)) - K''_{u^2}(s, t, x_0(t))]\, x(t)\, x'(t)\, dt \right| \leqslant$$

$$\leqslant \| x \| \left\{ \int_0^1 | [K''_{u^2}(s, t, x_0(t) + \theta\tau x'(t)) - \right.$$

$$\left. - K''_{u^2}(s, t, x_0(t))] \, x'(t) |^{\frac{p}{p-1}} dt \right\}^{\frac{p-1}{p}}. \qquad (4)$$

Si $x_0(t)$ et $x'(t)$ sont finies, l'intégrant de la dernière intégrale tend vers zéro avec $\tau$. Par ailleurs, en vertu de (2)

$$| [K''_{u^2}(s, t, x_0(t) + \theta\tau x'(t)) - K''_{u^2}(s, t, x_0(t))] \, x'(t) |^{\frac{p}{p-1}} \leqslant$$

$$\leqslant \{ [M | x_0(t) + \theta\tau x'(t) |^{p-2} + M | x_0(t) |^{p-2} + 2N] | x'(t) | \}^{\frac{p}{p-1}} \leqslant$$

$$\leqslant A_5 | x_0(t) |^{\frac{(p-2)p}{p-1}} | x'(t) |^{\frac{p}{p-1}} + A_6 | x'(t) |^p \; {}^{*)}.$$

Les produits $| x_0(t) |^{p - \frac{p}{p-1}} | x'(t) |^{\frac{p}{p-1}}$ et $| x'(t) |^p$ étant sommables, le passage à la limite est licite sous le signe somme dans (4). Il vient

$$\lim_{\tau \to 0} v_\tau(s) = v(s) \quad (0 \leqslant s \leqslant 1);$$

la limite étant manifestement uniforme en $x$, $\| x \| = 1$.

Les raisonnements précédents entraînent aussi

$$| v_\tau(s) | \leqslant \int_0^1 | K''_{u^2}(s, t, x_0(t) + \theta\tau x'(t)) \, x(t) \, x'(t) | \, dt \leqslant A_7,$$

de sorte que dans l'intégrale $\| v_\tau - v \|_{L^q} = \left\{ \int_0^1 | v_\tau(s) - v(s) |^q ds \right\}^{1/q}$ le passage à la limite est possible sous le signe somme, c'est-à-dire $\| v_\tau - v \| \to 0$ uniformément en $x$, $\| x \| = 1$.

C.q.f.d.

REMARQUE. Si l'on remplace la condition (2) par

$$| K''_{u^2}(s, t, u) | \leqslant M | u |^\lambda + N \quad (\lambda < p - 2) \qquad (5)$$

et si $p > 2$, on prouve que l'opérateur $P$, traité comme un opérateur de $\mathbf{L}^p$ dans $\mathbf{L}^\infty$, est deux fois dérivable en tout point $x_0 \in \mathbf{L}^p$, tel que $x_0$ soit une fonction bornée.

3.2. L'opérateur (1) peut être considéré comme un opérateur de $\mathbf{C}[0, 1]$ dans $\mathbf{C}[0, 1]$.

---

*) Il faut se servir de l'inégalité $| a + b |^\alpha \leqslant 2^\alpha (| a |^\alpha + | b |^\alpha)$ pour $\alpha > 0$.

Soit $x_0 \in C\,[0,\ 1]$ un élément fixe. Désignons par $\Omega$ la boule de $C\,[0,\ 1]$ de centre $x_0$ et de rayon $r$, et par $G$ l'ensemble des points $(s,\ t,\ u)$ de l'espace à trois dimensions, défini par les inégalités $0 \leqslant s,\ t \leqslant 1,\ |u - x_0(t)| \leqslant r$.

THÉORÈME 2. *Supposons que la fonction $K(s,\ t,\ u)$ est définie et continue dans $G$ avec ses dérivées $K'_u(s,\ t,\ u)$ et $K''_{u^2}(s,\ t,\ u)$. Alors l'opérateur $P$ défini par* (1) *envoie l'ensemble $\Omega$ dans* $C\,[0,\ 1]$, *est deux fois dérivable en chaque point intérieur $\bar{x} \in \Omega$, et $P'(\bar{x}) = U$ et $P''(\bar{x}) = B$ sont définies par les formules*

$$z = U(x), \quad z(s) = \int_0^1 K'_u(s,\ t,\ \bar{x}(t))\, x(t)\, dt,$$

$$v = B(x,\ x'), \quad v(s) = \int_0^1 K''_{u^2}(s,\ t,\ \bar{x}(t))\, x(t)\, x'(t)\, dt.$$

DÉMONSTRATION. Prenons un $x$ arbitraire de $\Omega$. La continuité de la fonction $y = P(x)$, définie par (1), résulte directement de faits élémentaires de la théorie des intégrales dépendant d'un paramètre, donc, $y \in C\,[0,\ 1]$.

Supposons maintenant que $\bar{x} \in \Omega$ et $x \in C\,[0,\ 1]$. Considérons l'élément

$$z_\tau = \frac{P(\bar{x} + \tau x) - P(\bar{x})}{\tau} - U(x).$$

On a

$$z_\tau(s) = \int_0^1 \Big[ \frac{K(s,\ t,\ \bar{x}(t) + \tau x(t)) - K(s,\ t,\ \bar{x}(t))}{\tau} - $$

$$- K'_u(s,\ t,\ \bar{x}(t))\, x(t) \Big]\, dt.$$

L'intégrant tend vers zéro avec $\tau$ et est uniformément borné en $x$ ($\|x\| = 1$) et en $s \in [0,\ 1]$ (on établit ce fait à l'aide de la formule des accroissements finis). Donc, $z_\tau \to 0$ uniformément en $x \in C\,[0,\ 1]$, $\|x\| = 1$. Ceci prouve la dérivabilité de $P$ en $\bar{x}$ et de plus

$$P'(\bar{x}) = U.$$

Par des raisonnements analogues on démontre que $P''(\bar{x}) = B$ et de plus que $P$ est deux fois dérivable.

3.3. Considérons un opérateur non linéaire lié à un système d'équations différentielles ordinaires de premier ordre:

$$\frac{dy_i}{ds} = f_i(s,\ y_k(s)) \quad (i,\ k = 1,\ 2,\ \ldots,\ n),$$

que nous écrirons pour la commodité sous la forme du système d'équations intégrales

$$y_i(s) = y_i^{(0)} + \int_0^s f_i(t, y_k(t))\,dt \quad (i, k = 1, 2, \ldots, n)$$

ou sous la forme vectorielle condensée

$$y(s) = y^{(0)} + \int_0^s f(t, y(t))\,dt,$$

où $y(s)$, $y^{(0)}$, $f(t, u)$, $u$ sont des vecteurs dont les composantes sont respectivement

$$(y_1(s), y_2(s), \ldots, y_n(s)), \quad (y_1^{(0)}, y_2^{(0)}, \ldots, y_n^{(0)}),$$

$$(f_1(t, u_k), \quad f_2(t, u_k), \ldots, f_n(t, u_k)), \quad (u_1, u_2, \ldots, u_n).$$

Considérons l'espace $C_n^\alpha$ ($\alpha > 0$) dont les éléments sont des fonctions vectorielles absolument continues sur $[0, \infty[$, telles que

$$\sup_t |y(t)e^{\alpha t}| < \infty, \quad \sup_t |y'(t)e^{\alpha t}| < \infty \quad (t \geqslant 0)\ *).$$

La norme est définie dans $C_n^\alpha$ par

$$\|y\| = \sup_t |y(t)e^{\alpha t}| + \sup_t |y'(t)e^{\alpha t}| \quad (y \in C_n^\alpha).$$

Considérons par ailleurs l'espace $D_n^\alpha$ des fonctions vectorielles absolument continues telles que

$$\sup_t |z(t)| < \infty, \quad \sup_t |z'(t)e^{\alpha t}| < \infty \quad (z \in D_n^\alpha).$$

Dans $D_n^\alpha$ la norme est définie par

$$\|z\| = \sup_t |z(t)| + \sup_t |z'(t)e^{\alpha t}| \quad (z \in D_n^\alpha).$$

Il est immédiat de vérifier que les espaces $C_n^\alpha$ et $D_n^\alpha$ sont des B-espaces.

Considérons l'opérateur $P$:

$$z = P(y), \tag{6}$$

$$z(s) = y(s) - \int_0^s f(t, y(t))\,dt. \tag{7}$$

---

*) Comme plus haut (cf. chap. XVI), nous désignons la longueur du vecteur $u$ par $|u|$ et la norme de la matrice $A$ considérée comme un opérateur de $l_n^2$ dans $l_n^2$ par $|A|$.

THÉORÈME 3. *Soit $f(t, u)$ une fonction vectorielle définie sur un ensemble $G$ d'un espace de dimension $n+1$;*

$$(t, u) \in G \quad \textit{signifie que} \quad t \geqslant 0, \quad |u| \leqslant re^{-\alpha t}.$$

*Ceci étant,*

1) $f(t, 0) = 0$;

2) *dans $G$ existe $f'_u(t, u)$* *) *bornée, continue en $u$ uniformément en $t \geqslant 0$.*

*Si l'on désigne par $\Omega$ la boule de l'espace $\mathrm{C}_n^\alpha$ de centre $0$ et de rayon $r$, alors l'opérateur* (6) *envoie $\Omega$ dans $\mathrm{D}_n^\alpha$, est dérivable en chaque point de $\Omega$, et de plus $P'(y_0) = U$, où $U$ est défini par*

$$z = U(y), \quad z(s) = y(s) - \int_0^1 f'_u(t, y_0(t))\, y(t)\, dt. \tag{8}$$

DÉMONSTRATION. Montrons tout d'abord que la fonction définie par (7) appartient à $\mathrm{D}_n^\alpha$ (à condition que $y \in \Omega$). En effet, comme $f(t, 0) = 0$, on a

$$z(s) = y(s) - \int_0^s [f(t, y(t)) - f(t, 0)]\, dt.$$

L'intégrant se majore avec la formule des accroissements finis de la manière suivante:

$$|f(t, y(t)) - f(t, 0)| \leqslant \sup_{0<\theta<1} |f'_u(t, \theta y(t))|\; |y(t)| \leqslant M \|y\| e^{-\alpha t}$$

$$(M = \sup |f'_u(t, u)|, \quad (t, u) \in G).$$

Cette majoration nous permet d'écrire

$$\sup_s |z(s)| \leqslant \sup_s |y(s)| + M\|y\| \sup_s \int_0^s e^{-\alpha t}\, dt \leqslant \|y\| + \frac{M}{\alpha}\|y\| < \infty.$$

Par ailleurs

$$z'(s) = y'(s) - f(s, y(s)),$$

et par des raisonnements analogues on s'assure que

$$\sup_s |z'(s)\, e^{\alpha s}| < \infty.$$

Donc, $z$ a un sens et appartient à l'espace $\mathrm{D}_n^\alpha$.

Vérifions maintenant que l'opérateur $U$, défini par la formule (8), est un opérateur linéaire continu de $\mathrm{C}_n^\alpha$ dans $\mathrm{D}_n^\alpha$.

---

*) On rappelle que la dérivée $f'_u(t, u)$ de la fonction vectorielle $f(t, \cdot)$ est la matrice $n$-carrée des dérivées $\left(\frac{\partial f_i}{\partial u_k}\right)$ (cf. 1.6).

En répétant les raisonnements précédents, on trouve pour $z = U(y)$

$$\sup_s |z(s)| + \sup_s |z'(s) e^{\alpha s}| =$$

$$= \sup_s \left| y(s) - \int_0^s f'_u(t, y_0(t))\, y(t)\, dt \right| +$$

$$+ \sup_s |[y'(s) - f'_u(s, y_0(s))\, y(s)]\, e^{\alpha s}| \leqslant$$

$$\leqslant \|y\| + \frac{M}{\alpha}\|y\| + (M+1)\|y\| = L\|y\|,$$

de sorte que $z = U(y) \in \mathbf{D}_n^\alpha$ et $\|z\| \leqslant L\|y\|$.

Comme dans les théorèmes précédents, considérons l'élément

$$v_\tau = \frac{P(y_0 + \tau y) - P(y_0)}{\tau} - U(y), \quad v_\tau(s) = \int_0^s J_\tau(t)\, dt$$

$$\left( J_\tau(t) = \frac{f(t, y_0(t) + \tau y(t)) - f(t, y_0(t))}{\tau} - f'_u(t, y_0(t))\, y(t) \right).$$

En appliquant la formule des accroissements finis à l'intégrant, on obtient

$$|J_\tau(t)| \leqslant \sup_{0<\theta<1} |f'_u(t, y_0(t) + \theta\tau y(t)) - f'_u(t, y_0(t))|\, |y(t)|.$$

D'où, compte tenu de la continuité uniforme de $f'_u(t, u)$, il suit que quel que soit $\varepsilon > 0$, pour $\tau$ assez petit, on a uniformément en $y$, $\|y\| = 1$,

$$|J_\tau(t)| \leqslant \varepsilon e^{-\alpha t},$$

par conséquent

$$\sup_s |v_\tau(s)| \leqslant \frac{\varepsilon}{\alpha}.$$

Pour les mêmes $\tau$ on aura de façon analogue

$$\sup_s |\dot{v}_\tau(s)\, e^{\alpha s}| \leqslant \varepsilon,$$

de sorte que

$$\|v_\tau\| \leqslant \varepsilon \left( \frac{1}{\alpha} + 1 \right).$$

Donc, $v_\tau \to 0$ pour $\tau \to 0$ uniformément en $y \in \mathbf{C}_n^\alpha$, $\|y\| = 1$, c.q.f.d.

## § 4. Théorème des fonctions implicites

**4.1.** Soient X et Y deux $B$-espaces. Formons l'ensemble $X \times Y$ de tous les couples $(x, y)$ $(x \in X, y \in Y)$. L'ensemble $X \times Y$ qui est linéarisé de façon naturelle, devient un $B$-espace si on le munit de la norme

$$\| (x, y) \| = \| x \| + \| y \|.$$

Ce $B$-espace est appelé *produit direct* des espaces X et Y (cf. IV.1.8, tome 1).

Si l'on identifie les couples de la forme $(x, 0_Y)$ $(x \in X$, $0_Y$ est le zéro de l'espace Y) avec les éléments de l'espace X, on peut admettre que X, et de façon analogue Y, sont des sous-espaces du produit direct $X \times Y$. Ceci étant, chaque élément $u = (x, y) \in X \times Y$ se représente de façon unique par

$$u = (x, y) = (x, 0_Y) + (0_X, y) = x + y.$$

Chaque opérateur linéaire $U \in B\,(X \times Y, Z)$ induit un couple $(U_X, U_Y)$ d'opérateurs linéaires $U_X \in B\,(X, Z)$, $U_Y \in B\,(Y, Z)$:

$$U_X(x) = U((x, 0_Y)), \quad U_Y(y) = U((0_X, y)).$$

Inversement, étant donné le couple $(U_X, U_Y)$, l'opérateur $U$

$$U((x, y)) = U_X(x) + U_Y(y)$$

est un opérateur linéaire continu de $X \times Y$ dans Z.

Soit $P$ un opérateur non linéaire de $\Omega \subset X \times Y$ dans un $B$-espace Z. Fixons $y_0 \in Y$ et considérons l'opérateur

$$P^{(y_0)} = P(\cdot, y_0)$$

de l'ensemble $\Omega^{(y_0)}$ des $x \in X$ tels que $(x, y_0) \in \Omega$ (la «coupure» de l'ensemble $\Omega$) dans Z. On définit de façon analogue l'opérateur $P^{(x_0)}$ $(x_0 \in X)$ de $\Omega^{(x_0)} \subset Y$ dans Z.

Si $x_0 \in X$ est un point intérieur de l'ensemble $\Omega^{(y_0)}$, on peut envisager la dérivée $P^{(y_0)\prime}(x_0)$. Cette dérivée est appelée *dérivée partielle par rapport à x* de l'opérateur $P$ en point $(x_0, y_0)$ et notée $P'_x(x_0, y_0)$. On définit de façon analogue $P'_y(x_0, y_0)$, dérivée partielle par rapport à $y$. De toute évidence

$$P'_x(x_0, y_0) \in B\,(X, Z), \quad P'_y(x_0, y_0) \in B\,(Y, Z).$$

Si l'opérateur $P$ admet au point $(x_0, y_0) \in \Omega$ la dérivée $P'(x_0, y_0) = U$, alors $P'_x(x_0, y_0) = U_X$ et $P'_y(x_0, y_0) = U_Y$; cette circonstance justifie les notations des dérivées partielles.

L'opérateur $P$ peut être traité comme une « fonction abstraite de deux variables», ce que nous traduirons en écrivant $P(x, y)$ au lieu de $P((x, y))$. De telles « fonctions» sont justiciables des

résultats fondamentaux de la théorie élémentaire des fonctions de plusieurs variables. Signalons l'un d'eux.

Supposons que les dérivées $P'_x$ et $P'_y$ existent dans un voisinage du point $(x_0, y_0) \in \Omega$. On a alors

$$\| P(x_0+\Delta x, y_0+\Delta y)-P(x_0, y_0)-P'_x(x_0, y_0)(\Delta x)- \\ -P'_y(x_0, y_0)(\Delta y)\| \leqslant \sup_{0<\theta, \theta_1<1} \| P'_x(x_0+\theta\Delta x, y_0+\theta_1\Delta y)- \\ -P'_x(x_0, y_0)\| \, \|\Delta x\| + \sup_{0<\theta, \theta_1<1} \| P'_y(x_0+\theta\Delta x, y_0+\theta_1\Delta y)- \\ -P'_y(x_0, y_0)\| \, \|\Delta y\|. \quad (1)$$

En effet,

$$A = \| P(x_0 + \Delta x, y_0 + \Delta y) - P(x_0, y_0) - P'_x(x_0, y_0)(\Delta x) - \\ - P'_y(x_0, y_0)(\Delta y)\| \leqslant \| P(x_0 + \Delta x, y_0 + \Delta y) - \\ - P(x_0, y_0 + \Delta y) - P'_x(x_0, y_0)(\Delta x)\| + \\ + \| P(x_0, y_0 + \Delta y) - P(x_0, y_0) - P'_y(x_0, y_0)(\Delta y)\|.$$

En appliquant à chaque terme la formule des accroissements finis (cf. (8), § 1), on obtient

$$A \leqslant \sup_{0<\theta<1} \| P'_x(x_0+\theta\Delta x, y_0+\Delta y)-P'_x(x_0, y_0)\| \, \|\Delta x\| + \\ + \sup_{0<\theta_1<1} \| P'_y(x_0, y_0+\theta_1\Delta y)- \\ -P'_y(x_0, y_0)\| \, \|\Delta y\| \leqslant \sup_{0<\theta, \theta_1<1} \| P'_x(x_0+\theta\Delta x, y_0+\theta_1\Delta y)- \\ -P'_x(x_0, y_0)\| \, \|\Delta x\| + \sup_{0<\theta, \theta_1<1} \| P'_y(x_0+\theta\Delta x, y_0+\theta_1\Delta y)- \\ -P'_y(x_0, y_0)\| \, \|\Delta y\|,$$

c.q.f.d.

4.2. Le principal résultat de ce paragraphe concerne l'existence de la solution de l'équation

$$P(x, y) = 0,$$

c'est-à-dire l'existence d'une fonction implicite, définie par cette équation. Comme en analyse élémentaire on a le

THÉORÈME 1. *Soit donné un opérateur $P$ de $\Omega$, voisinage d'un point $(x_0, y_0) \in X \times Y$, dans un espace $Z$, défini sur $\Omega$ et continu en $(x_0, y_0)$. Si*

1) $P(x_0, y_0) = 0$;
2) *dans $\Omega$ existe $P'_y$ continue en $(x_0, y_0)$,*
3) *l'opérateur $P'_y(x_0, y_0) \in B(Y, Z)$ admet un inverse continu*

$$\Gamma = [P'_y(x_0, y_0)]^{-1} \in B(Z, Y),$$

*alors il existe un opérateur $F$ défini sur un voisinage $G \subset \mathbf{X}$ du point $x_0$, de $G$ dans $\mathbf{Y}$, et possédant les propriétés suivantes:*

a) $P(x, F(x)) = 0 \quad (x \in G)$,

b) $F(x_0) = y_0$,

c) *$F$ est continu en $x_0$.*

*L'opérateur $F$ est défini de façon unique par* a), b), c) *en ce sens que si existe un opérateur $F_1$ doué des mêmes propriétés, on peut exhiber un* $\eta > 0$, *tel que pour* $\| x - x_0 \| < \eta$

$$F_1(x) = F(x).$$

Demonstration. Sans nuire à la généralité on peut admettre que $x_0 = 0_{\mathbf{X}}$, $y_0 = 0_{\mathbf{Y}}$. Prenons $x \in \mathbf{X}$ assez petit pour que $0_{\mathbf{Y}} \in \Omega^{(x)}$, où comme plus haut $\Omega^{(x)}$ représente l'ensemble des éléments $y \in \mathbf{Y}$, tels que $(x, y) \in \Omega$, et considérons sur $\Omega^{(x)}$ l'opérateur

$$Q^{(x)}(y) = y - \Gamma P(x, y).$$

Prouvons que pour tout nombre $\varepsilon > 0$ arbitrairement petit, on peut exhiber un $\delta > 0$ tel que pour $\| x \| < \delta$ l'opérateur $Q^{(x)}$ envoie la boule $\| y \| \leqslant \varepsilon$ dans elle-même. A cet effet, calculons et évaluons la dérivée de l'opérateur $Q^{(x)}$. D'après les propositions III et V de 1.2, on a

$$Q^{(x)\prime}(\bar{y})(y) = y - \Gamma P'_y(x, \bar{y})(y) = -\Gamma [P'_y(x, \bar{y}) - P'_y(0, 0)](y)$$

et par suite

$$\| Q^{(x)\prime}(\bar{y}) \| \leqslant \| \Gamma \| \, \| P'_y(x, \bar{y}) - P'_y(0, 0) \|.$$

L'opérateur $P'_{\mathbf{Y}}$ étant par hypothèse continu en $(0, 0)$, le second facteur peut être rendu arbitrairement petit, de sorte qu'en choisissant $\varepsilon$ et $\delta$ assez petits, on peut rendre

$$\| Q^{(x)\prime}(\bar{y}) \| \leqslant \alpha < 1 \quad (\| x \| < \delta, \quad \| \bar{y} \| \leqslant \varepsilon). \tag{2}$$

Majorons $Q^{(x)}(0)$. La condition 1) nous dit que

$$\| Q^{(x)}(0) \| = \| \Gamma P(x, 0) \| \leqslant \| \Gamma \| \, \| P(x, 0) - P(0, 0) \|,$$

donc, puisque l'opérateur $P$ est continu en $(0, 0)$, on peut en faisant décroître $\delta$ rendre le second membre de l'inégalité aussi petit que l'on veut. Nous admettrons si cela est nécessaire que $\delta$ est tel que

$$\| Q^{(x)}(0) \| \leqslant \varepsilon (1 - \alpha) \quad (\| x \| < \delta). \tag{3}$$

Désormais, il est aisé de démontrer la proposition énoncée plus haut à l'aide de la formule des accroissements finis. En effet, si $\| x \| < \delta$, $\| y \| < \varepsilon$, alors en vertu de (2) et (3)

$$\| Q^{(x)}(y) \| \leqslant \| Q^{(x)}(0) \| + \| Q^{(x)}(y) - Q^{(x)}(0) \| \leqslant$$
$$\leqslant \varepsilon (1 - \alpha) + \sup_{0<\theta<1} \| Q^{(x)\prime}(\theta y) \| \, \| y \| \leqslant \varepsilon (1 - \alpha) + \alpha\varepsilon = \varepsilon.$$

Ainsi, l'opérateur $Q^{(x)}$ envoie la boule fermée $\| y \| \leqslant \varepsilon$ dans elle-même, et de plus la dérivée $Q^{(x)\prime}$ vérifie la relation (2). On se trouve dans les conditions du théorème 1.1 et par suite la boule indiquée renferme l'unique point fixe $y^* = F(x)$ de cet opérateur:

$$y^* = y^* - \Gamma P(x, y^*),$$

c'est-à-dire

$$P(x, y^*) = 0.$$

L'opérateur $F$ est l'opérateur cherché. En effet, a) a déjà été vérifiée. b) découle de la condition 1), qui peut encore s'écrire $Q_0(0) = 0$, et de l'unicité du point fixe de l'opérateur $Q_0$. Enfin c) résulte de ce que $\varepsilon$ peut être choisi aussi petit que l'on veut, car il n'est pas minoré.

Prouvons l'unicité de $F$. La boule $\| y \| \leqslant \varepsilon$ contient l'unique point fixe de l'opérateur $Q^{(x)}$ ($\| x \| < \delta$), d'autre part, l'opérateur $F_1$ étant continu, pour $\eta$ assez petit on aura

$$\| F_1(x) \| = \| F_1(x) - F_1(0) \| \leqslant \varepsilon \quad (\| x \| < \eta),$$

de sorte que pour ces $x$ on aura $F_1(x) = F(x)$.

THÉORÈME 2. *Si dans le théorème 1 on admet que l'opérateur P est continu en tout point de $\Omega$, alors l'opérateur F est continu dans un voisinage du point $x_0$.*

DÉMONSTRATION. Dans le cas considéré, l'opérateur $Q^{(x)}$ est justiciable du théorème XVI.1.3. En effet, l'opérateur $Q^{(x)}$ ($\| x \| < \delta$) envoie la boule $\| y \| \leqslant \varepsilon$ dans elle-même, quant à l'inégalité (2), elle exprime que $Q^{(x)}$ est un opérateur de contraction (uniformément en $x \in X$, $\| x \| < \delta$). La continuité de $Q^{(x)}$ en le paramètre (ici c'est $x$) résulte de celle de l'opérateur $P$.

En appliquant le théorème indiqué, on obtient que $y^* = F(x)$ dépend continûment de $x$ pour $\| x \| < \delta$.

REMARQUE. Les théorèmes 1 et 2 revêtent un caractère local. Par ailleurs, ils ne fournissent pas de méthode efficace de majoration de $\varepsilon$ en fonction de $\delta$. Cette lacune sera comblée dans le chapitre suivant où l'utilisation de la dérivée seconde permet d'obtenir des résultats plus précis.

Nous pouvons déjà annoncer un résultat à l'aide du théorème 3 suivant.

4.3. THÉORÈME 3. *Supposons que les hypothèses du théorème 1 sont réunies et que, de plus, il existe dans $\Omega$ une dérivée partielle $P'_x$ continue au point $(x_0, y_0)$. Alors l'opérateur F est dérivable au point $x_0$ et de plus*

$$F'(x_0) = U,$$

*où*

$$U = -\Gamma P'_x(x_0, y_0) = -[P'_y(x_0, y_0)]^{-1} P'_x(x_0, y_0). \tag{4}$$

DÉMONSTRATION. On admettra comme dans la démonstration du théorème 1 que $x_0 = 0$, $y_0 = 0$.

D'après 1.1 il faut prouver que pour $\varepsilon > 0$ on peut exhiber $\delta > 0$ tel que pour tout $x \in X$, $\| x \| < \delta$, on ait

$$\| F(x) - F(0) - U(x) \| \leqslant \varepsilon \| x \|,$$

c'est-à-dire

$$\| F(x) - U(x) \| \leqslant \varepsilon \| x \|, \tag{5}$$

puisque $F(0) = 0$.

Mettons l'expression comprise sous le signe de la norme au premier membre de (5) sous la forme

$$F(x) - U(x) = y + \Gamma P'_x(0, 0)\, x = \Gamma [P'_x(0, 0)\, x + P'_y(0, 0)\, y],$$

où $F(x) = y$ et $U$ a été remplacé par son expression (4). Or,

$$P(x, y) = P(0, 0) = 0,$$

donc, en vertu de (1)

$$\begin{aligned} \| F(x) - U(x) \| &\leqslant \\ &\leqslant \| \Gamma \| \, \| P(x, y) - P(0, 0) - P'_x(0, 0)\, x - P'_y(0, 0)\, y \| \leqslant \\ &\leqslant \| \Gamma \| \, [ \sup_{0<\theta,\, \theta_1<1} \| P'_x(\theta x, \theta_1 y) - P'_x(0, 0) \| \cdot \| x \| + \\ &+ \sup_{0<\theta,\, \theta_1<1} \| P'_y(\theta x, \theta_1 y) - P'_y(0, 0) \| \, \| y \| ] \leqslant \eta [\| x \| + \| y \|], \end{aligned}$$

de plus les dérivés partielles $P'_x$ et $P'_y$ étant continues, la quantité $\eta$ peut être rendue aussi petite que l'on veut par le choix de $\delta$ assez petit. Donc,

$$\begin{aligned} \| F(x) - U(x) \| &\leqslant \eta [\| x \| + \| F(x) \|] \leqslant \\ &\leqslant \eta [\| x \| + \| U(x) \| + \| F(x) - U(x) \|]. \end{aligned}$$

Il s'ensuit que

$$\| F(x) - U(x) \| \leqslant \eta \frac{1 + \| U \|}{1 - \eta} \| x \|,$$

pourvu que $\eta$ soit assez petit. Donc, si $\eta$ est choisi tel que

$$\eta \frac{1 + \| U \|}{1 - \eta} \leqslant \varepsilon,$$

alors (5) sera réalisée pour tous les $x$ assez petits, c.q.f.d.

4.4. Nous allons achever ce paragraphe par un exemple d'application du théorème 1.

Considérons le système d'équations différentielles

$$y'_i(s) = f_i(s, y_k(s)) \quad (i, k = 1, 2, \ldots, n), \tag{6}$$

et mettons-le comme dans XVII.3.3 sous la forme

$$y(s) = y(0) + \int_0^s f(t, y(t))\, dt. \tag{7}$$

Supposons que $f(t, 0) = 0$ $(t \geqslant 0)$. Le système (6) (ou l'équation (7)) admet la solution triviale. Cette solution vérifie les conditions initiales $y(0) = 0$.

Il se pose la question, d'une grande importance pratique, de savoir sous quelles conditions l'équation (7) admet une solution définie sur $[0, \infty[$, vérifiant des conditions initiales assez petites, et surtout, de savoir quand la tendance des conditions initiales vers zéro entraîne celle de la solution vers zéro sur la section tout entière $[0, \infty[$. Voici des définitions rigoureuses.

Supposons que pour tout $\varepsilon > 0$ on puisse trouver un $\delta > 0$ tel que l'équation

$$y(s) = x + \int_0^s f(t, y(t))\, dt \tag{8}$$

admette une solution unique bornée $y^*$ dès que la norme du vecteur $x$ est $<\delta$, et de plus

$$\sup_{s \geqslant 0} |y^*(s)| \leqslant \varepsilon.$$

On dit alors que la solution triviale de l'équation (7) est *stable*. Si, en outre, $y^*(s) \underset{s \to \infty}{\longrightarrow} 0$, alors la solution triviale est dite *asymptotiquement stable*.

Avant d'aborder les conditions suffisantes de stabilité, voyons une notion auxiliaire.

Soit $T$ une matrice carrée d'ordre $n$. Formons la série

$$E + \frac{1}{1!}T + \frac{1}{2!}T^2 + \ldots + \frac{1}{n!}T^n + \ldots,$$

où $E$ est la matrice unité. Cette série converge quelle que soit la matrice $T$, puisque $|T^k| \leqslant |T|^k$ $(k = 0, 1, \ldots)$. Soit $e^T$ sa somme.

Si $x_0$ est un vecteur fixe, la fonction vectorielle

$$u(s) = e^{sT}x_0 \quad (s \geqslant 0)$$

vérifie l'équation différentielle

$$u'(s) = Tu(s), \tag{9}$$

ce dont on peut s'assurer exactement comme pour des fonctions réelles ordinaires. Cette circonstance est utilisée dans la démonstration du lemme suivant.

LEMME 1. *Soient* $\mu_1, \mu_2, \ldots$ *les valeurs propres de la matrice* $T$. *Les éléments de la matrice* $e^{sT}$ *sont de la forme* $\sum_k p_k(s)\, e^{s\mu_k}$, *où* $p_k$ *sont des polynômes de degré* $\leqslant n$.

*Donc, si* $\mu = \max_k \operatorname{Re}\mu_k$, *la norme de la matrice* $e^{sT}$ *admet la majoration suivante*:

$$| e^{sT} | \leqslant K_\eta e^{(\mu+\eta)s}, \tag{10}$$

*où* $\eta > 0$ *est arbitraire.*

DÉMONSTRATION. La première proposition résulte immédiatement de la théorie des systèmes d'équations différentielles linéaires à coefficients constants, si l'on tient compte du fait qu'en vertu de la remarque ci-dessus les colonnes de la matrice $e^{sT}$ forment un système linéairement indépendant complet de solutions de l'équation (9) qui en fait est un système d'équations différentielles linéaires à coefficients constants *).

D'où (10), puisque la norme de la matrice est majorée par la racine carrée de la somme des carrés de ses éléments.

Revenons maintenant à l'équation (7). Supposons comme dans 3.3 qu'une fonction vectorielle $f(t, u)$ est définie dans un domaine $G$ d'un espace de dimension $n+1$ ($t \geqslant 0$, $| u | \leqslant re^{-\alpha t}$, $\alpha > 0$), continue en $u$ pour $u = 0$ avec sa dérivée $f'_u(t, u)$. Supposons que la matrice $A = f'_u(t, 0)$ ne dépend pas de $t$ et désignons ses valeurs propres par $\lambda_1, \lambda_2, \ldots$

THÉORÈME 4. *Si*

$$\operatorname{Re}\lambda_k < 0 \quad (k = 1, 2, \ldots),$$

*alors la solution triviale de l'équation* (7) *est asymptotiquement stable. De façon plus précise, on a*

$$| y^*(s) | \leqslant L_x e^{-\alpha s} \quad (s \geqslant 0),$$

*où la constante* $L_x$ *dépend uniquement de* $x$ *et tend vers* 0 *avec* $x$, *et* $\alpha > 0$ *tel que* $\operatorname{Re}\lambda_k < -\alpha$ $(k = 1, 2, \ldots)$.

DÉMONSTRATION. Elle s'appuie sur le théorème 1. Posons dans ce théorème $\mathbf{X} = \mathbf{l}_n^2$, $\mathbf{Y} = \mathbf{C}_n^\alpha$, $\mathbf{Z} = \mathbf{D}_n^\alpha$, définissons l'opérateur $P$ comme suit :

$$z = P(x, y), \quad z(s) = y(s) - x - \int_0^s f(t, y(t))\, dt,$$

et prenons pour $\Omega$ l'ensemble des points $(x, y) \in \mathbf{X} \times \mathbf{Y}$, où $x$ est un élément quelconque de $\mathbf{X}$ et $\| y \| < r$.

*) Voir Stépanov [1], page 214.

15—01292

Le théorème 3.3 nous dit que les deux premières conditions du théorème 1 sont remplies. La vérification de la condition 3) est plus compliquée. Tout d'abord, le théorème 3.3 affirme que la dérivée $P'_Y(0, 0) = U$ existe et est de la forme

$$z = U(y), \quad z(s) = y(s) - \int_0^s Ay(t)\,dt,$$

donc, il faut prouver que l'opérateur $U$ admet un inverse continu. Les espaces $\mathbf{C}_n^\alpha$ et $\mathbf{D}_n^\alpha$ étant complets, il suffit de montrer que l'opérateur $U$ réalise une bijection entre $\mathbf{C}_n^\alpha$ et $\mathbf{D}_n^\alpha$ (cf. XII. 1.3), c'est-à-dire pour tout $\tilde{z} \in \mathbf{D}_n^\alpha$ existe un élément unique $\tilde{y} \in \mathbf{C}_n^\alpha$, tel que $\tilde{z} = U(\tilde{y})$. Autrement dit, il faut prouver que l'équation

$$\tilde{z}(s) = y(s) - \int_0^s Ay(t)\,dt \tag{11}$$

admet une solution unique dans $\mathbf{C}_n^\alpha$.

Il résulte de la théorie des équations différentielles que l'équation (11) admet toujours une solution unique de la forme

$$\tilde{y}(s) = \int_0^s e^{(s-t)A}\tilde{z}'(t)\,dt + e^{sA}z(0).$$

Donc, si l'on prend $\sigma$ tel que

$$\alpha < \sigma < -\max_k \operatorname{Re}\lambda_k,$$

en vertu de la majoration (10) on aura

$$|\tilde{y}(s)| \leqslant K\left[\int_0^s e^{-(s-t)\sigma}|\tilde{z}'(t)|\,dt + e^{-s\sigma}|\tilde{z}(0)|\right] \leqslant$$

$$\leqslant K\left[e^{s-\sigma}\int_0^s e^{(\sigma-\alpha)t}\|\tilde{z}\|\,dt + e^{-s\sigma}\|\tilde{z}\|\right] =$$

$$= K\|\tilde{z}\|e^{-s\sigma}\left[\frac{e^{(\sigma-\alpha)s}-1}{\sigma-\alpha}+1\right] =$$

$$= K\|\tilde{z}\|e^{-\alpha s}\left[\frac{1-e^{(\alpha-\sigma)s}}{\sigma-\alpha}+e^{(\alpha-\sigma)s}\right] \leqslant K_1\|\tilde{z}\|e^{-\alpha s},$$

de sorte que

$$\sup|\tilde{y}(s)e^{\alpha s}| \leqslant K_1\|\tilde{z}\| < \infty. \tag{12}$$

Comme

$$\tilde{y}'(s) = \tilde{z}'(x) - A\tilde{y}(s),$$

alors

$$\sup_s |\tilde{y}'(s)\, e^{\alpha s}| \leqslant \sup_s |\tilde{z}'(s)\, e^{\alpha s}| +$$
$$+ |A| \sup_s |\tilde{y}(s)\, \varepsilon^{\alpha s}| \leqslant (1 + K_1 |A|)\, \|\tilde{z}\| < \infty. \qquad (13)$$

La conjonction de (12) et (13) donne $\tilde{y} \in C_n^\alpha$, et les conditions du théorème 1 sont entièrement vérifiées.

Ce qui vient d'être prouvé nous permet donc d'affirmer que l'équation $P(x, y) = 0$, c'est-à-dire l'équation (8) admet une solution unique $y^* = F(x) \in C_n^\alpha$ si seulement le vecteur initial $x$ est assez petit. On a par ailleurs

$$|y^*(s)| \leqslant \|F(x)\|\, e^{-\alpha s},$$

et de plus $\|F(x)\| \to 0$ pour $x \to 0$, puisque $F$ est continu.

C.q.f.d.

CHAPITRE XVIII

# MÉTHODE DE NEWTON

Dans ce chapitre on développe en détail la théorie d'une méthode de résolution des équations fonctionnelles, connue dans le cas d'équations réelles sous le nom de méthode de Newton ou méthode des tangentes. Cette méthode et l'une de ses modifications sont parmi les rares méthodes utilisées en pratique pour la résolution des équations fonctionnelles non linéaires.

Cette méthode est d'une grande importance théorique, car elle permet de conclure à l'existence, l'unicité et la disposition de la solution sans chercher cette dernière, ce qui souvent est aussi profitable que la connaissance de cette solution.

Les résultats de ce chapitre sont dus essentiellement à L. Kantorovitch (cf. Kantorovitch [9]). L'exposé suit l'article de Kantorovitch [12].

## § 1. Equations de la forme $P(x) = 0$

**1.1.** Soit donné un opérateur $P$ d'un ouvert $\Omega$ d'un $B$-espace X dans un $B$-espace Y. Supposons que l'ensemble $\Omega$ contient un zéro de l'opérateur $P$, c'est-à-dire un élément $x^*$ tel que

$$P(x^*) = 0.$$

Soit $x_0$ un élément quelconque de $\Omega$. Si l'on suppose que l'opérateur $P$ admet une dérivée continue dans $\Omega$, on peut remplacer l'élément $P(x_0) = P(x_0) - P(x^*)$ par l'expression voisine $P'(x_0)(x_0 - x^*)$ et considérer, par suite, que la solution de l'équation

$$P'(x_0)(x_0 - x) = P(x_0)$$

sera proche de $x^*$. L'équation écrite est linéaire et sa solution qui est facile à trouver est

$$x_1 = x_0 - [P'(x_0)]^{-1}(P(x_0))$$

(on admet bien sûr l'existence de l'opérateur $[P'(x_0)]^{-1}$).

En poursuivant cette procédure, on obtient à partir de $x_0$, approximation initiale, la suite $\{x_n\}$

$$x_{n+1} = x_n - [P'(x_n)]^{-1}(P(x_n)) \quad (n = 0, 1, \ldots). \tag{1}$$

Chaque $x_n$ est une solution approchée de l'équation

$$P(x) = 0, \tag{2}$$

et cette solution est d'autant plus exacte que $n$ est grand.

La méthode de formation de la suite $\{x_n\}$ s'appelle *méthode de Newton* *).

Il est clair que la méthode de Newton ne passe pas toujours. D'abord, $x_n$ peut quitter l'ensemble $\Omega$ pour un certain $n$, et, ensuite, même si cela ne se produisait pas, $[P'(x_n)]^{-1}$ pourrait ne pas exister.

Si la suite $\{x_n\}$ converge vers la racine $x^*$, et $x_0$ est pris assez proche de $x^*$, alors les opérateurs $P'(x_n)$ et $P'(x_0)$ différeront peu, car $P'$ est continu. Ceci nous permet de remplacer les formules (1) par les formules

$$x'_{n+1} = x'_n - [P'(x_0)]^{-1}(P(x'_n)) \quad (n = 0, 1, \ldots; \quad x'_0 = x_0), \tag{3}$$

qui en général fournissent de moins bonnes approximations mais sont bien plus simples.

Le procédé de formation de la suite $\{x'_n\}$ sera appelé *méthode modifiée de Newton* **).

Plus bas, on étudie en détail les conditions de réalisation de la méthode de Newton (la méthode principale et la méthode modifiée) et sa convergence, c'est-à-dire la convergence des suites $\{x_n\}$ et $\{x'_n\}$ vers la solution de l'équation (2).

A noter que la méthode de Newton pour l'équation (2) est confondue avec la méthode ordinaire des approximations successives appliquée à l'équation

$$x = x - \Gamma(x)(P(x)) \quad (\Gamma(x) = [P'(x)]^{-1}), \tag{4}$$

qui équivaut visiblement à l'équation donnée.

De façon analogue, la méthode de Newton modifiée n'est autre que la méthode des approximations successives appliquée à l'équation

$$x = x - \Gamma(x_0)(P(x)). \tag{5}$$

Grâce à cette constatation nous commencerons par étudier la méthode ordinaire des approximations successives.

*) Si l'opérateur $P$ est une fonction réelle d'un argument réel, la formule (1) peut encore s'écrire

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)} \quad (n = 0, 1, \ldots),$$

ce qui nous conduit à la méthode de Newton proprement dite.

**) Ces deux méthodes sont souvent appelées méthodes de Newton-Kantorovitch.

1.2. Soit l'équation

$$x = S(x), \tag{6}$$

où $S$ est un opérateur défini dans la boule $\| x - x_0 \| < R$ d'un $B$-espace $\mathbf{X}$ ($x_0 \in \mathbf{X}$). Outre l'équation (6), considérons l'équation réelle

$$t = \varphi(t), \tag{7}$$

où la fonction $\varphi$ est définie sur l'intervalle $[t_0, t']$ ($t' = t_0 + r < < t_0 + R$). On dira que l'équation (7) (ou la fonction $\varphi$) *majore* l'équation (6) (ou l'opérateur $S$) si

$$\left.\begin{array}{l} 1)\ \| S(x_0) - x_0 \| \leqslant \varphi(t_0) - t_0; \\ 2)\ \| S'(x) \| \leqslant \varphi'(t) \text{ pour } \| x - x_0 \| \leqslant t - t_0. \end{array}\right\} \tag{8}$$

Théorème 1. *Supposons que l'opérateur $S$ possède une dérivée continue dans la boule fermée $\Omega_0$ ($\| x - x_0 \| \leqslant r$) et que la fonction $\varphi$ est dérivable sur l'intervalle $[t_0, t']$. Si l'équation* (7) *majore l'équation* (6) *et qu'elle possède une racine dans l'intervalle $[t_0, t']$, alors l'équation* (6) *admet aussi une solution $x^*$, limite de la suite $\{x_n\}$*

$$x_{n+1} = S(x_n) \quad (n = 0, 1, \ldots) \tag{9}$$

*des approximations successives d'origine $x_0$.*

*Ceci étant,*

$$\| x^* - x_0 \| \leqslant t^* - t_0, \tag{10}$$

*où $t^*$ représente la plus petite racine de l'équation* (7) *sur $[t_0, t_1]$.*

Démonstration. Montrons tout d'abord que les approximations successives pour l'équation (7)

$$t_{n+1} = \varphi(t_n) \quad (n = 0, 1, \ldots) \tag{11}$$

forment une suite convergente. Remarquons à cet effet que la condition 2) entraîne l'inégalité

$$\varphi'(t) \geqslant 0 \quad (t \in [t_0, t']),$$

si bien que la fonction $\varphi$ est croissante sur l'intervalle $[t_0, t']$. Il s'ensuit que $t_n$ a un sens pour tout $n$ et de plus que

$$t_n \leqslant \bar{t} \quad (n = 0, 1, \ldots), \tag{12}$$

où $\bar{t}$ désigne la racine de l'équation (7), dont l'existence est admise dans le théorème.

En effet, pour $n = 0$ l'inégalité (12) est évidente et si elle a été prouvée pour $n = k$, alors l'inégalité $t_k \leqslant \bar{t}$ et la monotonie de $\varphi$ entraînent que $\varphi(t_k) \leqslant \varphi(\bar{t})$, c'est-à-dire $t_{k+1} \leqslant \bar{t}$ et par récurrence l'inégalité est vraie pour tous les $n$.

En utilisant la monotonie de $\varphi$ on démontre aussi par récurrence celle de la suite $\{t_n\}$.

En effet, de $t_n \leqslant t_{n+1}$ il résulte que $t_{n+1} = \varphi(t_n) \leqslant \varphi(t_{n+1}) = t_{n+2}$, quant à l'inégalité $t_0 \leqslant t_1$, elle dérive de la condition 1).

On a donc établi l'existence de la limite $t^* = \lim_{n\to\infty} t_n$ qui, en vertu de (11) et de la continuité de $\varphi$, est racine de (7) en vertu de (12) la plus petite racine dans l'intervalle $[t_0, t']$.

Montrons à présent que tous les éléments de (9) ont un sens et forment une suite convergente.

Si l'on met (8) sous la forme

$$\| x_1 - x_0 \| \leqslant t_1 - t_0,$$

on constate que $x_1 \in \Omega_0$. Supposons qu'on ait déjà démontré que $x_1, x_2, \ldots, x_n \in \Omega_0$ et que

$$\| x_{k+1} - x_k \| \leqslant t_{k+1} - t_k \quad (k = 0, 1, \ldots, n-1). \tag{13}$$

Les résultats de XVII.1.7 nous disent que

$$x_{n+1} - x_n = S(x_n) - S(x_{n-1}) = \int_{x_{n-1}}^{x_n} S'(x)\,dx.$$

Si par $x$ et $t$ on désigne les points correspondants des intervalles $[x_{n-1}, x_n]$ et $[t_{n-1}, t_n]$, c'est-à-dire

$$x = x_{n-1} + \tau(x_n - x_{n-1}), \quad t = t_{n-1} + \tau(t_n - t_{n-1}) \quad (0 \leqslant \tau \leqslant 1),$$

alors en vertu de (13)

$$\| x - x_0 \| \leqslant \tau \| x_n - x_{n-1} \| + \| x_{n-1} - x_{n-2} \| + \ldots$$
$$\ldots + \| x_1 - x_0 \| \leqslant \tau(t_n - t_{n-1}) + (t_{n-1} - t_{n-2}) + \ldots$$
$$\ldots + (t_1 - t_0) = t - t_0.$$

Donc, d'après la condition 2)

$$\| S'(x) \| \leqslant \varphi'(t).$$

D'où, compte tenu de la remarque de XVII.1.7,

$$\| x_{n+1} - x_n \| = \left\| \int_{x_{n-1}}^{x_n} S'(x)\,dx \right\| \leqslant \int_{t_{n-1}}^{t_n} \varphi'(t)\,dt = \varphi(t_n) - \varphi(t_{n-1}) =$$
$$= t_{n+1} - t_n.$$

Donc, (13) a été prouvé pour $k = n$. On a aussi prouvé que $x_{n+1} \in \Omega_0$, puisque

$$\| x_{n+1} - x_0 \| \leqslant \| x_{n+1} - x_n \| + \| x_n - x_{n-1} \| + \ldots$$
$$\ldots + \| x_1 - x_0 \| \leqslant (t_{n+1} - t_n) + (t_n - t_{n-1}) + \ldots$$
$$\ldots + (t_1 - t_0) = t_{n+1} - t_0 \leqslant t' - t_0 = r.$$

Donc, on a établi par récurrence que pour tous les $k = 0, 1, \ldots$ on a $x_k \in \Omega_0$ et la majoration (13).

D'autre part, puisque en vertu de (13)

$$\| x_{n+p} - x_n \| \leqslant \| x_{n+p} - x_{n+p-1} \| + \ldots + \| x_{n+1} - x_n \| \leqslant$$
$$\leqslant (t_{n+p} - t_{n+p-1}) + \ldots + (t_{n+1} - t_n) = t_{n+p} - t_n, \quad (14)$$

alors la suite $\{x_n\}$ est de Cauchy et par conséquent admet une limite que l'on notera $x^* = \lim\limits_{n \to \infty} x_n$. En passant à la limite dans (9) et en tenant compte de la continuité de $S$, on obtient

$$x^* = S(x^*),$$

c'est-à-dire $x^*$ est solution de l'équation (6).

La majoration (10) découle de (14) si l'on y fait $n = 0$ et que l'on passe à la limite pour $p \to \infty$.

REMARQUE 1. L'inégalité

$$\| x^* - x_n \| \leqslant t^* - t_n \quad (n = 0, 1, \ldots) \quad (15)$$

qui résulte de (14) par passage à la limite pour $p \to \infty$ donne la majoration de la vitesse de convergence de la suite $\{x_n\}$ vers $x^*$.

REMARQUE 2. Pour démontrer le théorème on ne s'est pas entièrement servi de la condition 2). Il fallait que l'inégalité $\| S'(x) \| \leqslant \leqslant \varphi'(t)$ soit vérifiée seulement en les points correspondants des intervalles $[x_{n-1}, x_n]$ et $[t_{n-1}, t_n]$ $(n = 1, 2, \ldots)$.

L'équation (6) peut admettre dans la boule $\Omega_0$ des racines autres que $x^*$, même si la solution $t^*$ de l'équation majorante (7) est unique dans l'intervalle $[t_0, t']$. Cependant on a le

THÉORÈME 2. *Supposons que sont réalisées les conditions du théorème précédent et que de plus*

$$\varphi(t') \leqslant t'.$$

*Si l'équation* (7) *possède dans l'intervalle* $[t_0, t']$ *une solution unique, alors la boule* $\Omega_0$ *contient l'unique racine* $x^*$ *de l'équation* (6) *vers laquelle converge le processus des approximations successives d'origine* $\tilde{x}_0 \in \Omega_0$.

DÉMONSTRATION. Appliquons à l'équation (7) la méthode des approximations successives en prenant pour approximation initiale $\tilde{t}_0 = t'$. On démontre exactement comme dans le théorème 1 que la suite

$$\tilde{t}_{n+1} = \varphi(\tilde{t}_n) \quad (n = 0, 1, \ldots)$$

est monotone décroissante et minorée $(\tilde{t}_n \geqslant t^*)$, donc admet une limite égale à $\tilde{t}$ et de plus $\tilde{t}$ en tant que racine de l'équation (7) doit être confondue avec $t^*$ $(\tilde{t} = t^*)$.

Prouvons maintenant que le processus des approximations successives pour l'équation (6) converge, quelle que soit son origine $\tilde{x}_0 \in \Omega_0$, vers la solution de l'équation (6).

Les approximations successives sont de la forme

$$\tilde{x}_{n+1}=S(\tilde{x}_n) \quad (n=0,\ 1,\ \ldots).$$

On a

$$\tilde{x}_1-x_1=S(\tilde{x}_0)-S(x_0)=\int_{x_0}^{\tilde{x}_0} S'(x)\,dx,$$

d'où, grâce à la remarque de XVII.1.7,

$$\|\tilde{x}_1-x_1\|\leqslant\int_{t_0}^{\tilde{t}_0}\varphi'(t)\,dt=\varphi(\tilde{t}_0)-\varphi(t_0)=\tilde{t}_1-t_1.$$

Il est clair que

$$\|\tilde{x}_1-x_0\|\leqslant\|\tilde{x}_1-x_1\|+\|x_1-x_0\|\leqslant(\tilde{t}_1-t_1)+(t_1-t_0)=\tilde{t}_1-t_1\leqslant r,$$

donc $\tilde{x}_1\in\Omega_0$.

Supposons qu'on ait déjà prouvé que

$$\tilde{x}_k\in\Omega_0,\quad \|\tilde{x}_k-x_k\|\leqslant\tilde{t}_k-t_k \quad (k=1,\ 2,\ \ldots,\ n). \tag{16}$$

Il vient

$$\tilde{x}_{n+1}-x_{n+1}=S(\tilde{x}_n)-S(x_n)=\int_{x_n}^{\tilde{x}_n} S'(x)\,dx.$$

Mais si $x$ et $t$, les points correspondants des intervalles $[x_n,\ \tilde{x}_n]$ et $[t_n,\ \tilde{t}_n]$, sont

$$x=x_n+\tau(\tilde{x}_n-x_n),\quad t=t_n+\tau(\tilde{t}_n-t_n)\quad (0\leqslant\tau\leqslant 1),$$

alors

$$\|x-x_0\|\leqslant\tau\,\|\tilde{x}_n-x_n\|+\|x_n-x_{n-1}\|+\ldots+\|x_1-x_0\|\leqslant$$
$$\leqslant\tau(\tilde{t}_n-t_n)+(t_n-t_{n-1})+\ldots+(t_1-t_0)=t-t_0$$

et pour les $x$ et $t$ indiqués on aura $\|S'(x)\|\leqslant\varphi'(t)$ et

$$\|\tilde{x}_{n+1}-x_{n+1}\|\leqslant\int_{t_n}^{\tilde{t}_n}\varphi'(t)\,dt=\varphi(\tilde{t}_n)-\varphi(t_n)=\tilde{t}_{n+1}-t_{n+1}.$$

Or

$$\|\tilde{x}_{n+1}-x_0\|\leqslant\|\tilde{x}_{n+1}-x_{n+1}\|+\|x_{n+1}-x_0\|\leqslant$$
$$\leqslant(\tilde{t}_{n+1}-t_{n+1})+(t_{n+1}-t_0)=\tilde{t}_{n+1}-t_0\leqslant r,$$

donc $\tilde{x}_{n+1}\in\Omega_0$.

On établit par récurrence que (16) est valable pour tous les $k = 1, 2, \ldots$

Les suites $\{t_n\}$ et $\{\tilde{t}_n\}$ admettant une limite commune (égale à $t^*$), la convergence de la suite $\{x_n\}$, qui résulte de (16), entraîne celle de la suite $\{\tilde{x}_n\}$ et l'égalité

$$\lim_{n\to\infty}\tilde{x}_n=\lim_{n\to\infty}x_n=x^*. \tag{17}$$

On a donc prouvé que le processus des approximations successives converge vers $x^*$ quelle que soit l'approximation initiale $\tilde{x}_0 \in \Omega_0$. Il s'ensuit que l'équation (6) admet une solution unique. En effet, si $\tilde{x} \in \Omega_0$ est racine de cette équation, alors en prenant $\tilde{x}_0 = \tilde{x}$ on obtient manifestement

$$\tilde{x}_n=\tilde{x},\quad \forall n=1,\ 2,\ \ldots$$

et en vertu de (17) $\tilde{x} = x^*$, c.q.f.d.

**1.3.** Etudions la méthode de Newton pour l'équation (2).

Considérons avec (2) l'équation réelle

$$\psi(t) = 0. \tag{18}$$

On admettra que l'opérateur $P$ est défini dans la boule $\Omega$ ($\| x - x_0 \| < R$) de l'espace X et admet une dérivée seconde continue dans la boule fermée $\Omega_0$ ($\| x - x_0 \| \leqslant r$). On admettra que la fonction $\psi$ est deux fois continûment dérivable sur l'intervalle $[t_0, t']$ ($t' = t_0 + r$).

Grâce au théorème 1 on obtient aisément le résultat suivant sur la convergence de la méthode modifiée de Newton.

THÉORÈME 3. *Soient réalisées les conditions*:

1) *il existe un opérateur linéaire continu* $\Gamma_0 = [P'(x_0)]^{-1}$;

2) $c_0 = -\dfrac{1}{\psi'(t_0)} > 0$;

3) $\|\Gamma_0(P(x_0))\|\leqslant c_0\psi(t_0)$;

4) $\|\Gamma_0 P''(x)\|\leqslant c_0\psi''(t)$, *si* $\|x-x_0\|\leqslant t-t_0\leqslant r$;

5) *l'équation* (18) *possède une racine* $\bar{t}$ *dans l'intervalle* $[t_0, t']$.

*Alors le processus modifié de Newton pour les équations* (2) *et* (18), *d'origine respectivement* $x_0$ *et* $t_0$, *converge vers les solutions* $x^*$ *et* $t^*$ *de ces équations, et, de plus*

$$\| x^* - x_0 \| \leqslant t^* - t_0. \tag{19}$$

Démonstration. On a signalé plus haut que la méthode modifiée de Newton pour l'équation (2) est équivalente à la méthode des approximations successives pour l'équation (5), c'est-à-dire pour l'équation

$$x = S(x) \quad (S(x) = x - \Gamma_0(P(x)). \tag{20}$$

De façon analogue, la méthode modifiée de Newton pour l'équation (18) peut être remplacée par le processus des approximations successives pour l'équation

$$t = \varphi(t) \quad (\varphi(t) = t + c_0\psi(t)). \tag{21}$$

Prouvons que les conditions du théorème 1 sont réalisées pour les équations (20) et (21).

On a

$$S(x_0) - x_0 = -\Gamma_0(P(x_0))$$

et de façon analogue

$$\varphi(t_0) - t_0 = c_0\psi(t_0).$$

Donc, d'après la condition 3)

$$\| S(x_0) - x_0 \| \leqslant \varphi(t_0) - t_0.$$

Une dérivation (cf. XVII.1.2) nous donne

$$S'(x) = I - \Gamma_0 P'(x), \quad S''(x) = -\Gamma_0 P''(x),$$

donc

$$S'(x) = S'(x) - S'(x_0) = \int_{x_0}^{x} S''(x)\,dx = -\int_{x_0}^{x} \Gamma_0 P''(x)\,dx,$$

d'où, en vertu de la remarque de XVII. 1.7 et de la condition 4)

$$\| S'(x) \| \ll \int_{t_0}^{t} c_0\psi''(\tau)\,d\tau = c_0\psi'(t) - c_0\psi'(t_0) = 1 + c_0\psi''(t_0) = \varphi'(t),$$

pourvu que $\| x - x_0 \| \leqslant t - t_0$.

Donc, l'équation (21) majore l'équation (20). On peut donc leur appliquer le théorème 1 et obtenir le résultat voulu.

Remarque. Comme dans le théorème 1, la vitesse de convergence de la suite $\{x_n'\}$ (cf. (3)) vers $x^*$ est majorée par

$$\| x^* - x_n' \| \leqslant t^* - t_n' \quad (n = 0, 1, \ldots; \; x_0' = x_0; \quad t_0' = t_0, \tag{22}$$

où $t_n'$ sont les approximations successives de la méthode modifiée de Newton pour l'équation (18).

L'application du théorème 2 aux équations (20) et (21) permet d'établir immédiatement l'unicité de la solution de l'équation (2).

THÉORÈME 4. *Soient réalisées les conditions du théorème* 3 *et, de plus*

$$\psi(t') \leqslant 0. \tag{23}$$

*Si l'équation* (18) *admet une racine unique dans l'intervalle* $[t_0, t']$, *alors l'équation* (2) *admet une racine unique dans la boule* $\Omega_0$.

Pour en faire la preuve il suffit de remarquer que (23) entraîne

$$\varphi(t') = t' + c_0\psi(t') \leqslant t',$$

c'est-à-dire est réalisée la condition subsidiaire du théorème 2.

REMARQUE. $t^*$ étant la plus petite racine de l'équation (18), on peut se servir du théorème d'unicité si l'on pose $t' = t^*$. On peut donc affirmer que l'équation (2) possède une solution unique en tous les cas dans la boule

$$\| x - x_0 \| \leqslant t^* - t_0.$$

1.4. On établit la convergence du processus principal de Newton grâce au théorème 3 de convergence du processus modifié.

THÉORÈME 5. *Soient remplies les conditions du théorème* 3. *Le processus principal de Newton pour l'équation* (2), *commencé en* $x_0$, *est réalisable et conduit à une suite* $\{x_n\}$ *qui converge vers la racine* $x^*$ *de l'équation* (2).

DÉMONSTRATION. Le premier pas étant le même dans le processus principal et le processus modifié de Newton, $x_1$ a un sens et $x_1 \in \Omega_0$. Vérifions que les conditions du théorème 3 ne sont pas violées par la substitution de $x_1$ à $x_0$ et de $t_1$ à $t_0$. Conservons les notations du théorème 3 et considérons l'opérateur

$$I - \Gamma_0 P'(x_1) = -\Gamma_0 (P'(x_1) - P'(x_0)) = -\int_{x_0}^{x_1} \Gamma_0 P''(x)\, dx.$$

La remarque de XVII.1.7 nous donne

$$\| I - \Gamma_0 P'(x_1) \| \leqslant \int_{t_0}^{t_1} c_0 \psi''(t)\, dt = 1 + c_0 \varphi'(t_1) = q.$$

La fonction $\psi(t)$ ne peut atteindre son minimum à gauche du point $t^*$, puisque $\psi''(t) \geqslant 0$ sur l'intervalle $[t_0, t']$ tout entier et $\psi(t_0) \geqslant 0$. Comme $t_1 \leqslant t^*$ et $\psi'(t_0) < 0$, il vient $\psi'(t_1) < 0$ *) et par suite $q < 1$.

---

*) Si $\psi'(t_1) = 0$, alors $t_1 = t^*$ et $\psi(t_1) = 0$, et par suite

$$0 = \psi(t_1) - \psi(t_0) - \psi'(t_0)(t_1 - t_0) = \int_{t_0}^{t_1} \psi''(t)(t_1 - t)\, dt,$$

ce qui n'est possible que pour $\psi''(t) = 0$ sur l'intervalle $[t_0, t_1]$, ce qui nous conduit à l'égalité $\psi'(t_0) = \psi'(t_1) = 0$.

Le théorème de Banach (V.4.5, tome 1) affirme l'existence d'un opérateur linéaire continu

$$U=[\Gamma_0 P'(x_1)]^{-1}\in B(\mathrm{X},\ \mathrm{X}),$$

et de plus

$$\|U\|\leqslant\frac{1}{1-q}=\frac{\psi'(t_0)}{\psi'(t_1)}=\frac{c_1}{c_0}\quad\left(c_1=-\frac{1}{\psi'(t_1)}\right). \tag{24}$$

Il existe donc l'opérateur linéaire continu

$$\Gamma_1=[P'(x_1)]^{-1}=U\Gamma_0.$$

Les conditions 1) et 2) du théorème 3 ont bien lieu.
Montrons maintenant que

$$\|\Gamma_0(P(x_1))\|\leqslant c_0\psi(t_1). \tag{25}$$

La formule de Taylor (XVII.2.5) nous donne

$$\Gamma_0(P(x_1))=\Gamma_0(P(x_0))+\Gamma_0 P'(x_0)\ (x_1-x_0)+$$
$$+\int_{x_0}^{x_1}\Gamma_0 P''(x)(x_1-x,\cdot)\,dx=(x_0-x_1)+(x_1-x_0)+$$
$$+\int_{x_0}^{x_1}\Gamma_0 P''(x)(x_1-x,\cdot)\,dx=\int_{x_0}^{x_1}\Gamma_0 P''(x)(x_1-x,\cdot)\,dx.$$

De façon analogue

$$c_0\psi(t_1)=c_0\int_{t_0}^{t_1}\psi''(t)(t_1-t)\,dt.$$

Pour les points correspondants $x$ et $t$ des intervalles $[x_0, x_1]$ et $[t_0, t_1]$, on a

$$\|\Gamma_0 P''(x)(x_1-x,\ \cdot)\|\leqslant\|\Gamma_0 P''(x)\|\,\|x_1-x\|\leqslant c_0\psi''(t)(t_1-t),$$

donc

$$\left\|\int_{x_0}^{x_1}\Gamma_0 P''(x)(x_1-x,\ \cdot)\,dx\right\|\leqslant\int_{t_0}^{t_1}c_0\psi''(t)(t_1-t)\,dt,$$

ce qui équivaut à (25).

Les relations (25) et (24) nous donnent

$$\|\Gamma_1(P(x_1))\|=\|U\Gamma_0(P(x_1))\|\leqslant\|U\|\,\|\Gamma_0(P(x_1))\|\leqslant$$
$$\leqslant\frac{c_1}{c_0}c_0\psi(t_1)=c_1\psi(t_1),$$

c'est-à-dire la condition 3) est aussi remplie.

La condition 4) se vérifie de façon analogue. Signalons tout d'abord que si $\| x-x_1 \| \leqslant t-t_1$, alors *a fortiori* $\|x-x_0\| \leqslant t-t_0$. Donc,

$$\| \Gamma_1 P''(x) \| = \| U \Gamma_0 P''(x) \| \leqslant \| U \| \, \| \Gamma_0 P''(x) \| \leqslant \frac{c_1}{c_0} c_0 \psi''(t) = c_1 \psi''(t).$$

Enfin, la condition 5) n'est pas violée, car la racine $\bar{t}$ est comprise dans l'intervalle $[t^*, t']$, donc dans l'intervalle plus large $[t_1, t']$.

On montre par des raisonnements analogues que les conditions du théorème 3 ne sont pas violées par la substitution de $x_2$ et $t_2$ à $x_1$ et $t_1$, etc.

En définitive, tous les $x_n$ ont un sens, c'est-à-dire on a démontré que le processus de Newton est réalisable.

La suite $\{t_n\}$ est manifestement croissante et bornée, donc elle admet une limite $\tilde{t}$, quant à l'inégalité

$$\| x_{n+1} - x_n \| \leqslant t_{n+1} - t_n \quad (n=0, 1, \ldots)$$

elle nous dit que la suite $\{x_n\}$ admet également une limite $\tilde{x} = \lim\limits_{n \to \infty} x_n$.

De (1) l'on déduit la relation

$$P(x_n) + P'(x_n)(x_{n+1} - x_n) = 0 \quad (n=0, 1, \ldots). \tag{26}$$

Pour $x$ quelconque de $\Omega_0$, la formule des accroissements finis nous donne

$$\| \Gamma_0 [P'(x) - P'(x_0)] \| \leqslant$$
$$\leqslant \| x - x_0 \| \sup_{0<\theta<1} \| \Gamma_0 P''(x_0 + \theta(x - x_0)) \| \leqslant r \max_{[t_0, t']} c_0 \psi''(t),$$

d'où il suit que $\| P'(x_n) \|$ sont bornées dans leur ensemble. Donc, un passage à la limite dans (26) nous donne

$$P(\tilde{x}) = 0,$$

c'est-à-dire $\tilde{x}$ est racine de l'équation (2).

De façon analogue, $\tilde{t}$ est racine de l'équation (18) et de plus $\tilde{t} = t^*$, car $\tilde{t} \leqslant t^*$ et $t^*$ est la plus petite racine.

La remarque suivant le théorème 4 nous apprend que $\tilde{x} = x^*$, puisque manifestement $\| \tilde{x} - x_0 \| \leqslant \tilde{t} - t_0 = t^* - t_0$.

Ceci achève la démonstration du théorème.

Remarque. La vitesse de convergence de la suite $\{x_n\}$ vers $x^*$ admet la majoration

$$\| x^* - x_n \| \leqslant t^* - t_n \quad (n = 0, 1, \ldots). \tag{27}$$

Ceci découle de l'inégalité (22), puisque $x_n$ et $t_n$ peuvent être traités comme les premières approximations de la méthode de Newton modifiée d'origine $x_{n-1}$ et $t_{n-1}$ respectivement.

1.5. L'application directe du théorème 3 ou 5 présente généralement des difficultés en raison de la présence de la fonction indéfinie $\psi$ dans leurs énoncés. Le théorème suivant revêt de ce fait une signification particulière.

THÉORÈME 6. *On suppose toujours que l'opérateur $P$ est défini sur $\Omega$ et admet une dérivée seconde continue sur $\Omega_0$, et de plus*

1) *existe l'opérateur linéaire continu* $\Gamma_0 = [P'(x_0)]^{-1}$;
2) $\| \Gamma_0 (P(x_0)) \| \leqslant \eta$;
3) $\| \Gamma_0 P''(x) \| \leqslant K \ (x \in \Omega_0)$.

*Si*

$$h = K\eta \leqslant \frac{1}{2} \tag{28}$$

*et*

$$r \geqslant r_0 = \frac{1-\sqrt{1-2h}}{h}\eta, \tag{29}$$

*alors l'équation* (2) *admet une solution $x^*$ vers laquelle converge le processus de Newton (le principal et le modifié) et de plus*

$$\| x^* - x_0 \| \leqslant r_0. \tag{30}$$

*Si, par ailleurs, pour* $h < \frac{1}{2}$

$$r < r_1 = \frac{1+\sqrt{1-2h}}{h}\eta \tag{31}$$

*et pour* $h = \frac{1}{2}$

$$r \leqslant r_1, \tag{32}$$

*alors la solution $x^*$ est unique dans la boule $\Omega_0$.*

*La vitesse de convergence du processus principal admet la majoration*

$$\| x^* - x_n \| \leqslant \frac{1}{2^n}(2h)^{2^n}\frac{\eta}{h} \quad (n = 0, 1, \ldots), \tag{33}$$

*celle du processus modifié, la majoration*

$$\| x^* - x_n \| \leqslant \frac{\eta}{h}\left(1-\sqrt{1-2h}\right)^{n+1}, \quad h < \frac{1}{2}, \quad (n = 0, 1, \ldots). \tag{34}$$

DÉMONSTRATION. Considérons dans l'intervalle $[0, r]$ la fonction réelle

$$\psi(t) = Kt^2 - 2t + 2\eta = Kt^2 - 2t + \frac{2h}{K} = \frac{h}{\eta}t^2 - 2t + 2\eta.$$

Vérifions que l'opérateur $P$ et la fonction $\psi$ remplissent toutes les conditions du théorème 3 (donc, du théorème 5). En effet, les conditions 1) à 4) sont évidentes. D'autre part, les racines de l'équa-

tion

$$\psi(t) = 0 \tag{35}$$

sont

$$r_0 = \frac{1-\sqrt{1-2h}}{h}\eta, \quad r_1 = \frac{1+\sqrt{1-2h}}{h}\eta,$$

ces racines sont réelles d'après (28), quant à (29) elle nous dit que la plus petite racine, c'est-à-dire $r_0$, est contenue dans l'intervalle $[0, r]$. Comme $t^* = r_0$, l'inégalité (30) n'est autre que l'inégalité (19) du théorème 3.

L'unicité de la solution est une conséquence du théorème 4, car, quel que soit $h \leqslant 1/2$, le théorème nous dit que $\psi(r) \leqslant 0$ et la racine de l'équation (35) est unique dans l'intervalle $[0, r]$.

Etablissons maintenant les majorations (33) et (34) de la vitesse de convergence du processus de Newton. D'après les remarques suivant les théorèmes 3 et 5, il suffit de considérer les suites réelles $\{t_n\}$ et $\{t'_n\}$ des approximations successives respectivement des processus principal et modifié pour l'équation (35).

Commençons par la méthode principale. Posons

$$c_n = -\frac{1}{\psi'(t_n)}, \quad \eta_n = c_n\psi(t_n),$$

$$K_n = c_n\psi''(t_n) = 2Kc_n, \quad h_n = K_n\eta_n$$

et exprimons $\eta_n$, $K_n$ et $h_n$ en fonction de ces grandeurs, mais avec un indice inférieur d'une unité. Pour cela, on remarquera que

$$t_{k+1} - t_k = -\frac{\psi(t_k)}{\psi'(t_k)} = \eta_k \quad (k = 0, 1, \ldots). \tag{36}$$

Donc, la formule de Taylor nous donne

$$\eta_n = c_n\psi(t_n) = c_n\psi(t_{n-1} + \eta_{n-1}) =$$
$$= c_n\left[\frac{1}{2}\psi''(t_{n-1})\eta_{n-1}^2 + \psi'(t_{n-1})\eta_{n-1} + \psi(t_{n-1})\right] =$$
$$= c_n\left[K\eta_{n-1}^2 - \frac{\eta_{n-1}}{c_{n-1}} + \frac{\eta_{n-1}}{c_{n-1}}\right] = c_nK\eta_{n-1}^2 =$$
$$= \frac{1}{2}\frac{c_n}{c_{n-1}}2Kc_{n-1}\eta_{n-1}^2 = \frac{1}{2}\frac{c_n}{c_{n-1}}K_{n-1}\eta_{n-1}^2.$$

Or

$$\frac{c_{n-1}}{c_n} = \frac{\psi'(t_n)}{\psi'(t_{n-1})} = \frac{\psi'(t_{n-1}) + \psi''(t_{n-1})\eta_{n-1}}{\psi'(t_{n-1})} =$$
$$= 1 - K_{n-1}\eta_{n-1} = 1 - h_{n-1}, \tag{37}$$

donc

$$\eta_n = \frac{1}{2}\frac{K_n\eta_{n-1}^2}{1-h_{n-1}} = \frac{\eta_{n-1}}{2}\frac{h_{n-1}}{1-h_{n-1}}. \tag{38}$$

De façon analogue, on a grâce à (37)

$$K_n = 2c_n K = 2Kc_{n-1}\frac{c_n}{c_{n-1}} = \frac{K_{n-1}}{1-h_{n-1}}.$$

D'où

$$h_n = K_n\eta_n = \frac{1}{2}\,\frac{K_{n-1}\eta_{n-1}h_{n-1}}{(1-h_{n-1})^2} = \frac{1}{2}\left[\frac{h_{n-1}}{1-h_{n-1}}\right]^2. \tag{39}$$

De (38) et (39) et compte tenu de $h_n \leqslant 1/2$, il vient

$$\eta_n \leqslant h_{n-1}\eta_{n-1}, \qquad h_n \leqslant 2h_{n-1}^2 \quad (n = 1,\ 2,\ \ldots). \tag{40}$$

Donc

$$h_n \leqslant \frac{1}{2}\,[2h_0]^{2^n} = \frac{1}{2}\,[2h]^{2^n}$$

et

$$\eta_n \leqslant h_{n-1}\eta_{n-1} \leqslant h_{n-1}h_{n-2}\eta_{n-2} \leqslant \ldots$$
$$\ldots \leqslant h_{n-1}h_{n-2}\ldots h_0\eta_0 \leqslant \frac{1}{2^n}\,[2h]^{2^n-1}\eta,$$

d'où, grâce à (36), l'on déduit

$$t^* - t_n = (t_{n+1}-t_n) + (t_{n+2}-t_{n+1}) + \ldots = \eta_n + \eta_{n+1} + \ldots \leqslant$$
$$\leqslant \frac{1}{2^n}\,[2h]^{2^n-1}\eta\left\{1 + \frac{1}{2}\,[2h]^{2^n} + \frac{1}{2^2}\,[2h]^{2^{n+1}} + \ldots\right\} \leqslant$$
$$\leqslant \frac{2}{2^n}\,[2h]^{2^n-1}\eta = \frac{1}{2^n}\,[2h]^{2^n}\,\frac{\eta}{h}.$$

La majoration cherchée (35) découle désormais de (27).

Majorons maintenant l'erreur de la méthode modifiée. On suppose que $h < 1/2$. Si $\varphi$ est la fonction utilisée dans la démonstration du théorème 3, on peut écrire

$$t^* - t'_n = \varphi(t^*) - \varphi(t'_{n-1}) = \varphi'(\tilde{t}_n)\,(t^* - t'_{n-1})$$
$$\left(\tilde{t}_n = \frac{t^* + t'_{n-1}}{2}\right).$$

Or

$$\varphi'(t) = 1 + c_0\psi'(t) = Kt,$$

si bien que

$$\varphi'(\tilde{t}_n) = K\tilde{t}_n \leqslant Kt^* = 1 - \sqrt{1-2h}.$$

Donc

$$t^* - t'_n \leqslant [1 - \sqrt{1-2h}]\,(t^* - t'_{n-1}).$$

En reprenant ce raisonnement pour majorer $t^* - t'_{n-1}$ et en le poursuivant, on obtient en définitive

$$t^* - t'_n \leqslant [1 - \sqrt{1-2h}]^n\,(t^* - t'_0) = \frac{\eta}{h}\,[1 - \sqrt{1-2h}]^{n+1}.$$

En portant ceci dans (22), on obtient

$$\| x^* - x'_n \| \leqslant \frac{\eta}{h} [1 - \sqrt{1-2h}]^{n+1},$$

c'est-à-dire l'inégalité (34). C.q.f.d.

REMARQUE 1. Les conditions 2) et 3) du théorème peuvent être remplacées par:

1') $\| \Gamma_0 \| \leqslant B'$;
2') $\| P(x_0) \| \leqslant \eta'$;
3') $\| P''(x) \| \leqslant K'$ $(x \in \Omega_0)$.

Ceci étant

$$h = K'B'^2\eta',$$

$$r_0 = \frac{1-\sqrt{1-2h}}{h} B'\eta',$$

$$r_1 = \frac{1+\sqrt{1-2h}}{h} B'\eta'.$$

REMARQUE 2. Les conditions (28), (29), (31) (ou (32)) sont exactes en ce sens que si l'une des deux premières est violée l'équation (35) ne possède pas de solutions, et si l'une des deux dernières l'est, la solution n'est pas unique *).

## § 2. Corollaires du théorème de convergence de la méthode de Newton

Dans ce paragraphe, on indique quelques corollaires du théorème de convergence de la méthode de Newton (théorème 1.6). Comme dans le paragraphe précédent, on considère l'équation

$$P(x) = 0, \tag{1}$$

l'opérateur $P$ étant justiciable des anciennes hypothèses (quant au domaine de définition, des propriétés des dérivées, etc.) et notations.

2.1. Nous commençons par généraliser le théorème 1.6 au cas où l'on n'exige pas l'existence de l'opérateur inverse $\Gamma_0 = [P'(x_0)]^{-1}$, mais celle d'un opérateur proche de lui.

THÉORÈME 1. *Supposons qu'existe un opérateur linéaire* $\Gamma \in B(\mathrm{Y}, \mathrm{X})$ *possédant un inverse continu et que sont réalisées les conditions*:

1) $\| \Gamma(P(x_0)) \| \leqslant \overline{\eta}$;
2) $\| \Gamma P'(x_0) - I \| \leqslant \delta$;
3) $\| \Gamma P''(x) \| \leqslant \overline{K}$ $(x \in \Omega_0)$.

---

*) Si la condition (32) n'est pas réalisée, l'équation (30) admet une solution unique. Mais l'équation $\psi(t) - \varepsilon = 0$ $(\varepsilon > 0)$ admet déjà deux solutions sur l'intervalle $[0, r]$.

*Si*

$$\bar{h} = \frac{\bar{K}\bar{\eta}}{(1-\delta)^2} \leqslant \frac{1}{2}, \quad \delta < 1 \tag{2}$$

*et*

$$r \geqslant \bar{r}_0 = \frac{1-\sqrt{1-2\bar{h}}}{\bar{h}} \frac{\bar{\eta}}{1-\delta}, \tag{3}$$

*alors l'équation* (1) *possède une solution* $x^* \in \Omega_0$, *unique si*

$$\begin{matrix} r < \bar{r}_1 & \text{pour} & \bar{h} < 1/2, \\ r \leqslant \bar{r}_1 & \text{pour} & \bar{h} = 1/2 \end{matrix} \quad \left(\bar{r}_1 = \frac{1+\sqrt{1-2\bar{h}}}{\bar{h}} \frac{\bar{\eta}}{1-\delta}\right).$$

*Si de plus* $\{x_n\}$ *est la suite des approximations successives du processus principal de Newton, alors*

$$\| x^* - x_n \| \leqslant \frac{1}{2^n} [2\bar{h}]^{2^n} \frac{\bar{\eta}}{\bar{h}(1-\delta)} \quad (n = 0, 1, \ldots).$$

DÉMONSTRATION. Assurons-nous que les conditions du théorème 1.6 sont réalisées. La deuxième condition entraîne l'existence de l'opérateur linéaire continu

$$U = [\Gamma P'(x_0)]^{-1},$$

et de plus

$$\| U \| \leqslant \frac{1}{1-\delta}.$$

Il s'ensuit qu'existe l'opérateur linéaire continu

$$\Gamma_0 = [P'(x_0)]^{-1} = U\Gamma.$$

En outre

$$\| \Gamma_0 (P(x_0)) \| \leqslant \| U \| \, \| \Gamma (P(x_0)) \| \leqslant \frac{\bar{\eta}}{1-\delta},$$

$$\| \Gamma_0 P''(x) \| \leqslant \| U \| \, \| \Gamma P''(x) \| \leqslant \frac{\bar{K}}{1-\delta} \quad (x \in \Omega_0).$$

Reste à remplacer $\eta$ par $\frac{\bar{\eta}}{1-\delta}$ et $K$ par $\frac{\bar{K}}{1-\delta}$ dans le théorème 1.6.

2.2. Considérons un processus d'approximations successives identique au processus modifié et remplaçons dans ce dernier l'opérateur $\Gamma_0 = [P'(x_0)]^{-1}$ par un opérateur $\Gamma$ proche de lui. En d'autres termes, construisons la suite $\{\tilde{x}_n\}$

$$\tilde{x}_{n+1} = \tilde{x}_n - \Gamma(P(\tilde{x}_n)) \quad (n = 0, 1, \ldots; \ \tilde{x}_0 = x_0). \tag{4}$$

THÉORÈME 2. *Dans les conditions du théorème* 1, *le processus* (4) *est réalisable, c'est-à-dire* $\tilde{x}_n \in \Omega_0$ $(n = 0, 1, \ldots)$ *et*

$$\tilde{x}_n \to x^*.$$

DÉMONSTRATION. Pour démontrer la convergence du processus modifié, on se servira, comme dans le théorème 1.3, du théorème 1.1 dans lequel on posera

$$S(x) = x - \Gamma(P(x)) \quad (x \in \Omega_0)$$

et

$$\varphi(t) = \frac{1}{2}\overline{K}t^2 + \delta t + \overline{\eta} = t + \psi(t)$$

$$\left(\psi(t) = \frac{1}{2}Kt^2 - (1-\delta)t + \overline{\eta}, \quad t \in [t_0, t'], \quad t_0 = 0, \quad t' = r\right).$$

On a

$$\| S(x_0) - x_0 \| = \| -\Gamma(P(x_0)) \| \leqslant \overline{\eta} = \varphi(t_0) - t_0.$$

D'autre part, si

$$\| x - x_0 \| \leqslant t \leqslant r,$$

alors

$$\| S'(x) \| = \| I - \Gamma P'(x) \| \leqslant$$
$$\leqslant \| I - \Gamma P'(x_0) \| + \| \Gamma(P'(x_0) - P'(x)) \| \leqslant$$
$$\leqslant \delta + \left\| \int_{x_0}^{x} \Gamma P''(y)\, dy \right\| \leqslant \delta + \int_0^t \overline{K}\, d\tau = \delta + \overline{K}t = \varphi'(t).$$

Donc, l'équation

$$x = S(x) \tag{5}$$

est majorée par l'équation $t = \varphi(t)$ dont la solution est

$$t^* = \frac{1-\sqrt{1-2\overline{h}}}{\overline{h}} \frac{\overline{\eta}}{1-\delta} = \overline{r}_0 \leqslant r.$$

Par conséquent, l'équation (5) possède une solution qui doit être confondue avec $x^*$ (puisque l'équation (1) admet une solution unique $x^*$ dans la boule $\| x - x_0 \| \leqslant \overline{r}_0$).

REMARQUE. La vitesse de convergence de la suite $\{\tilde{x}_n\}$ vers $x^*$ peut être majorée exactement comme nous l'avons fait pour le processus modifié dans 1.4. On laisse notamment au lecteur le soin de prouver que pour $h < \frac{1}{2}$

$$\| x^* - \tilde{x}_n \| \leqslant \frac{1}{\overline{h}} \left[1 - (1-\delta)\sqrt{1-2\overline{h}}\right]^{n+1} \frac{\overline{\eta}}{(1-\delta)^2}$$
$$(n = 0, 1, \ldots).$$

2.3. On se place dans les conditions du théorème 1.6, avec $h < \frac{1}{2}$, et l'on prouve que le processus de Newton est s t a b l e, c'est-à-dire la convergence subsiste si pour approximation initiale on prend non pas $x_0$ mais un élément quelconque $x_0' \in \Omega_0$ assez proche de lui.

THÉORÈME 3. *Supposons que sont réalisées les conditions du théorème 1.6 avec les constantes* $\eta$, $K$ *et* $h = K\eta < 1/2$. *Si de plus*

$$\| x' - x_0 \| \leqslant \varepsilon \left( \varepsilon = \frac{1-2h}{4K} = \frac{1-2K\eta}{4K} \right), \tag{6}$$

*alors les processus principal et modifié de Newton convergent si pour approximation initiale on prend* $x_0'$.

DÉMONSTRATION. On se sert du théorème 1 dans lequel on remplace $\Gamma$ par $\Gamma_0 = [P'(x_0)]^{-1}$ et $x_0$ par $x_0'$. La condition 1) devient

$$\| \Gamma_0 (P(x_0')) \| =$$

$$= \left\| \Gamma_0 \left[ P(x_0) + P'(x_0)(x_0' - x_0) + \int_{x_0}^{x_0'} P''(x)(x_0' - x, \cdot)\, dx \right] \right\| \leqslant$$

$$\leqslant \eta + \varepsilon + K\frac{\varepsilon^2}{2},$$

donc pour $\overline{\eta}$ on doit prendre $\eta + \varepsilon + K\frac{\varepsilon^2}{2}$.

$$\| \Gamma_0 P'(x_0') - I \| = \| \Gamma_0 [P'(x_0') - P'(x_0)] \| = \left\| \int_{x_0}^{x_0'} \Gamma_0 P''(x)\, dx \right\|.$$

Donc $\delta = K\varepsilon$.

Enfin $\overline{K} = K$.

La quantité $\overline{h}$ vaut

$$\overline{h} = \frac{\overline{K}\,\overline{\eta}}{(1-\delta)^2} = \frac{K\left(\eta + \varepsilon + \frac{K}{2}\varepsilon^2\right)}{(1-K\varepsilon)^2} = \frac{1}{2}\,\frac{K^2\varepsilon^2 + 2K\varepsilon + 2K\eta}{K^2\varepsilon^2 - 2K\varepsilon + 1} = \frac{1}{2}.$$

C.q.f.d.

REMARQUE. Si $h \geqslant 4\sqrt{2} - 5{,}5 \approx 0{,}16$, la boule $\| x - x_0 \| \leqslant \varepsilon$ est contenue dans la boule $\| x - x_0 \| \leqslant r_0$, donc dans la boule $\Omega_0$.

Si $h < 4\sqrt{2} - 5{,}5$, la condition (6) n'assure plus d'une façon générale l'appartenance de $x_0'$ à $\Omega_0$, il faut donc l'exiger accessoirement.

2.4. Parmi les données qui ont participé à la majoration de $\| x^* - x_0 \|$, c'est-à-dire de $r_0$, figure l'opérateur $\Gamma_0$. La connaissance de ce dernier permet de mieux localiser la solution. En effet, si l'on connaît $\Gamma_0$, on peut trouver l'approximation suivante $x_1$ à l'aide de la méthode de Newton et lui appliquer le théorème 1.

THÉORÈME 4. *Supposons qu'existe l'opérateur linéaire continu* $\Gamma_0 = [P'(x_0)]^{-1}$ *et que sont réalisées les conditions*:

1) $\| \Gamma_0 (P(x_0)) \| \leqslant \eta$;

2) $\| \Gamma_0 (P(x_1)) \| \leqslant \eta_1$ $(x_1 = x_0 - \Gamma_0 (P(x_0)))$;

3) $\| \Gamma_0 P''(x) \| \leqslant K \quad (x \in \Omega_0)$;

4) $h_1 = \frac{K\eta_1}{(1-K\eta)^2} \leqslant \frac{1}{2}$.

*Alors l'équation* (1) *admet la solution* $x^*$ *et de plus*

$$\| x^* - x_1 \| \leqslant \frac{1-\sqrt{1-2h_1}}{h_1} \frac{\eta_1}{1-K\eta}. \tag{7}$$

Démonstration. Vérifions les conditions du théorème 1. En posant $\Gamma = \Gamma_0$ et en prenant $x_1$ au lieu de $x_0$, on obtient

$$\| \Gamma_0 (P(x_1)) \| \leqslant \eta_1,$$

$$\| \Gamma_0 P'(x_1) - I \| \leqslant \| \Gamma_0 [P'(x_1) - P'(x_0)] \| =$$

$$= \left\| \int_{x_0}^{x_1} \Gamma_0 P''(x)\, dx \right\| \leqslant K \| x_1 - x_0 \| \leqslant K\eta.$$

Pour $\bar{\eta}$, $\delta$ et $\bar{h}$ il faut prendre respectivement $\eta_1$, $K\eta$ et $h_1$.

Le théorème 1 nous dit que la solution $x^*$ est contenue dans la boule $\| x - x_1 \| \leqslant \bar{r}_0$. Comme

$$\bar{r}_0 = \frac{1-\sqrt{1-2\bar{h}}}{\bar{h}} \frac{\bar{\eta}}{1-\delta} = \frac{1-\sqrt{1-2h_1}}{h_1} \frac{\eta_1}{1-K\eta},$$

il vient (7).

Remarque 1. Il peut arriver que $h_1 \leqslant 1/2$, bien que $h = K\eta > 1/2$. Dans ce cas, le théorème permet de conclure à l'existence de la solution, bien que le théorème 1.6 ne soit pas appliquable.

Remarque 2. Les conditions 1) à 4) n'impliquent pas l'inclusion de la boule $\| x - x_1 \| < \bar{r}_0$ dans la boule $\Omega_0$. Il faut donc admettre que le rayon de la boule $\Omega_0$ est assez grand pour que cette inclusion ait lieu, par exemple

$$r \geqslant \eta + \bar{r}_0 = \eta + \frac{1-\sqrt{1-2h_1}}{h_1} \frac{\eta_1}{1-K\eta}.$$

2.5. Si dans les conditions de la remarque 1 qui suit le théorème 1.6, la majoration de la norme de l'opérateur $[P'(x)]^{-1}$ est connue non seulement en $x_0$ mais sur le domaine $\Omega_0$ tout entier, on peut affaiblir la condition imposée à $h$ et exiger que $h < 2$ au lieu de $h < 1/2$ (cf. Myssovskikh [1]).

Théorème 5 (Myssovskikh). *Soient réalisées les conditions*:

1) $\| P(x_0) \| \leqslant \eta$;

2) *pour* $x \in \Omega_0$ *existe l'opérateur linéaire continu* $\Gamma(x) = [P'(x)]^{-1}$ *et de plus*

$$\| \Gamma(x) \| \leqslant B \quad (x \in \Omega_0);$$

3) $\| P''(x) \| \leqslant K \quad (x \in \Omega_0)$.

*Si*

$$h = B^2K\eta < 2$$

*et*

$$r > r' = B\eta \sum_{k=0}^{\infty} \left(\frac{h}{2}\right)^{2^k-1},$$

*alors l'équation* (1) *admet une solution* $x^* \in \Omega_0$ *vers laquelle converge le processus principal de Newton d'origine* $x_0$, *la vitesse de convergence étant majorée par*

$$\|x^* - x_n\| \leqslant B\eta \frac{\left(\frac{h}{2}\right)^{2^n-1}}{1-\left(\frac{h}{2}\right)^{2^n}}$$

$$(x_{n+1} = x_n - \Gamma(x_n)(P(x_n)); \quad n = 0, 1, \ldots). \tag{9}$$

DÉMONSTRATION. Prouvons tout d'abord que le processus de Newton est réalisable, c'est-à-dire que $x_n \in \Omega_0$ $(n = 0, 1, \ldots)$. On a

$$\|x_1 - x_0\| = \|-\Gamma(x_0)(P(x_0))\| \leqslant B\eta, \tag{10}$$

$$\|P(x_1)\| = \left\| P(x_0) + P'(x_0)(x_1 - x_0) + \int_{x_0}^{x_1} P''(x)(x_1 - x, \cdot)\, dx \right\| \leqslant$$

$$\leqslant \int_0^{B\eta} K(B\eta - t)\, dt = \frac{KB^2\eta^2}{2} = \frac{h\eta}{2} = \eta_1. \tag{11}$$

Il résulte de (10) que $x_1 \in \Omega_0$.

Calculons la quantité $h$ correspondant au point $x_1$:

$$h_1 = B^2K\eta_1 = \frac{B^2K\eta h}{2} = \frac{h^2}{2}. \tag{12}$$

Supposons qu'on ait déjà prouvé que $x_n \in \Omega_0$. En désignant par $h_k$ et $\eta_k$ $(k = 0, 1, \ldots)$ les quantités $h$ et $\eta$ correspondant au point $x_k$, on obtient par analogie avec (10), (11) et (12)

$$\|x_{n+1} - x_n\| \leqslant B\eta_n,$$

$$\eta_{k+1} = \frac{h_k\eta_k}{2} \quad (k = 0, 1, \ldots, n-1), \tag{13}$$

$$h_k = 2\left(\frac{h_{k-1}}{2}\right)^2 = \ldots = 2\left(\frac{h_0}{2}\right)^{2^k} = 2\left(\frac{h}{2}\right)^{2^k} \quad (k = 0, 1, \ldots, n).$$

Les deux dernières relations nous donnent

$$\eta_n = \frac{h_{n-1}\eta_{n-1}}{2} = \frac{h_{n-1}h_{n-2}\eta_{n-2}}{4} = \ldots = \frac{h_{n-1}h_{n-2}\ldots h_0}{2^n}\eta_0 =$$

$$= \left(\frac{h}{2}\right)^{2^n-1}\eta.$$

En portant ceci dans (13), on obtient

$$\| x_{n+1} - x_n \| \leqslant B\eta \left(\frac{h}{2}\right)^{2^n - 1}. \tag{14}$$

Par analogie, il vient pour $\| x_{k+1} - x_k \|$ $(k = 0, 1, \ldots, n-1)$

$$\| x_{n+1} - x_0 \| \leqslant \sum_{k=0}^{n} \| x_{k+1} - x_k \| \leqslant B\eta \sum_{k=0}^{n} \left(\frac{h}{2}\right)^{2^k - 1} \leqslant r' \leqslant r$$

et par suite $x_{n+1} \in \Omega_0$.

En vertu de (14) on a de même

$$\| x_{n+p} - x_n \| \leqslant \sum_{k=n}^{n+p-1} \| x_{k+1} - x_k \| \leqslant \sum_{k=n}^{n+p-1} B\eta \left(\frac{h}{2}\right)^{2^k - 1} \xrightarrow[n \to \infty]{} 0, \tag{15}$$

si bien que la suite $\{x_n\}$ est de Cauchy, donc existe $x^* = \lim\limits_{n \to \infty} x_n$. Il est clair que $x^*$ est solution de l'équation (1).

En faisant tendre $p \to \infty$ dans (15), on obtient

$$\| x^* - x_n \| \leqslant B\eta \sum_{k=n}^{\infty} \left(\frac{h}{2}\right)^{2^k - 1} \leqslant$$

$$\leqslant B\eta \left(\frac{h}{2}\right)^{2^n - 1} \sum_{k=0}^{\infty} \left(\frac{h}{2}\right)^{2^n k} = B\eta \frac{(n/2)^{2^n - 1}}{1 - (h/2)^{2^n}},$$

ce qui nous donne la majoration (9) de la vitesse de convergence du processus.

**2.6.** Il est souvent possible de remplacer l'équation donnée (1) par une équation proche plus simple, généralement non linéaire. Etudions les conditions de solubilité de l'équation donnée, sachant la solution de l'équation approchée.

Supposons que l'équation (1) est de la forme

$$P(x) \equiv \pi(x) + R(x) = 0. \tag{16}$$

Supposons que $x_0$ est solution de l'équation simplifiée

$$\pi(x_0) = 0.$$

Si

1) $\| [\pi'(x_0)]^{-1} R(x_0) \| \leqslant \eta$;
2) $\| [\pi'(x_0)]^{-1} R'(x_0) \| \leqslant \alpha < 1$;
3) $\| [\pi'(x_0)]^{-1} \pi''(x) \| \leqslant K$, $\| [\pi'(x_0)]^{-1} R''(x) \| \leqslant L$ $(x \in \Omega_0)$

et

$$h = \frac{\eta (K+L)}{(1-\alpha)^2} \leqslant \frac{1}{2}, \tag{17}$$

$$r \geqslant r_0 = \frac{1-\sqrt{1-2h}}{h} \frac{\eta}{1-\alpha}, \tag{18}$$

alors l'équation (16) admet une solution $x^* \in \Omega_0$.

Cette proposition découle directement du théorème 1 si l'on y pose $\Gamma = [\pi'(x_0)]^{-1}$. Le théorème 1 nous permet encore de déterminer le domaine d'unicité de la solution $x^*$.

Cette situation est un cas particulier d'une situation plus générale où le premier membre de l'équation (1) dépend d'un paramètre numérique ou autre et où, sachant une solution de l'équation pour une valeur du paramètre, on demande d'établir l'existence d'une solution pour des valeurs du paramètre voisines de l'initiale. Nous nous bornerons au cas où le paramètre figure linéairement dans l'équation. Plus exactement, nous considérons l'équation

$$P(x, \mu) \equiv \pi(x) + \mu R(x) = 0, \tag{19}$$

où $\mu$ est un opérateur linéaire dans l'espace $\mathbf{Y}$ ($\mu \in B(\mathbf{Y}, \mathbf{Y})$). En particulier, $\mu$ peut être un facteur numérique.

Théorème 6. *Soient réalisées les conditions suivantes:*

1) $P(x_0, 0) = \pi(x_0) = 0$;

2) *existe l'opérateur linéaire continu* $\Gamma_0 = [\pi'(x_0)]^{-1}$ *et* $\| \Gamma_0 \| \leqslant B$;

3) $\| R(x_0) \| \leqslant \eta$, $\| R'(x_0) \| \leqslant \alpha$;

4) $\| \pi''(x) \| \leqslant K$, $\| R''(x) \| \leqslant L$ $(x \in \Omega_0)$.

*Si* $\mu$ *est tel que*

$$h_\mu = \frac{B^2 \eta (K + L \| \mu \|) \| \mu \|}{(1-\alpha B \| \mu \|)^2} \leqslant \frac{1}{2}, \quad \alpha B \| \mu \| < 1, \tag{20}$$

*alors l'équation* (19) *admet la solution* $x^* \in \Omega_0$ *pourvu que* $\Omega_0$ *possède un rayon* $r$ *assez grand.*

Demonstration. Elle découle directement de ce qui a été dit à propos de l'équation (16).

Les domaines de disposition et d'unicité de la solution $x^*(\mu)$ peuvent être trouvés à l'aide du théorème 1. Ce théorème nous dit que $x^*(\mu)$ est la limite des processus principal et modifié de Newton, d'origine $x_0$. De plus, le théorème 2 nous apprend que $x^*(\mu)$ est la limite de la suite $\{\tilde{x}_n(\mu)\}$:

$$\tilde{x}_{n+1}(\mu) = \tilde{x}_n(\mu) - [\pi'(x_0)]^{-1} (P(\tilde{x}_n(\mu), \mu)) \\ (n = 0, 1, \ldots; \tilde{x}_0(\mu) = x_0). \tag{21}$$

Le dernier processus présente l'avantage de faire intervenir l'opérateur inverse correspondant au point initial $x_0$ et à la valeur initiale $\mu = 0$ du paramètre.

REMARQUE 1. Pour chaque $\mu$, l'équation majorante de (19) sera l'équation réelle

$$\left(\frac{K}{2}t^2-\frac{1}{B}t\right)+\|\mu\|\left(\frac{L}{2}t^2+\alpha t+\eta\right)=0,$$

quant à la condition (20) elle exprime le fait que les racines de cette équation sont réelles.

REMARQUE 2. Des raisonnements plus compliqués nous permettent d'étudier le cas général où le paramètre $\mu$ figure non linéairement dans l'équation $P(x, \mu) = 0$.

A noter que dans certains cas, on peut élargir le domaine des valeurs admissibles du paramètre en appliquant le théorème 1 non pas à $x_0$ et $\Gamma_0$, mais à des valeurs $x$ et $\Gamma$ qui sont solutions du système d'équations

$$P(x, \mu) \equiv 0, \quad \Gamma P'(x, \mu) = I$$

aux « premières» puissances près du paramètre $\mu$.

Limitons-nous au cas où le paramètre est numérique et l'équation approchée $\pi(x) = 0$, linéaire, c'est-à-dire l'équation (19) est de la forme

$$P(x, \mu) \equiv U(x - x_0) + \mu R(x) = 0, \tag{22}$$

où $U = (\pi'(x_0))$ est un opérateur linéaire continu.

Dans ce cas les solutions, à $\mu^1$ près, sont de la forme

$$x_1 = x_0 - \mu\Gamma_0(R(x_0)), \quad \Gamma_1 = \Gamma_0 - \mu\Gamma_0 R'(x_0)\Gamma_0.$$

Il faut donc majorer les quantités

1) $\|\Gamma_1(P(x_1, \mu))\| = \|[\Gamma_0 - \mu\Gamma_0 R'(x_0)\Gamma_0][-\mu R(x_0) +$
$+ \mu R(x_0 - \mu\Gamma_0 R(x_0))]\| \leqslant \eta_1;$

2) $\|\Gamma_1 P'(x_1, \mu) - I\| = \|[\Gamma_0 - \mu\Gamma_0 R'(x_0)\Gamma_0][U +$
$+ \mu R'(x_0 - \mu\Gamma_0 R(x_0))] - I\| \leqslant \delta_1;$

3) $\|[\Gamma_0 - \mu\Gamma_0 R'(x_0)\Gamma_0] R''(x)\| \leqslant L_1 \quad (x \in \Omega_0).$

L'équation (22) admet une solution si

$$\frac{L_1\eta_1|\mu|}{(1-\delta_1)^2} \leqslant \frac{1}{2}.$$

2.7. L'existence de la solution et la convergence du processus correspondant dans le domaine de variation du paramètre permettent d'établir le caractère de la dépendance de la solution $x^*(\mu)$ par rapport à $\mu$. Par exemple, comme $\tilde{x}_0(\mu) = x_0$ dépend continûment de $\mu$, et $\pi(x)$ et $R(x)$ sont des fonctions continues de $x$, il vient que tous

les $\tilde{x}_n$ ($\mu$) de (21) seront des fonctions continues du paramètre, donc $x$ ($\mu$) dépendra continûment de $\mu$.

Sous certaines conditions, la solution $x^*$ ($\mu$) dépend analytiquement du paramètre.

On dira qu'une fonction $x$ ($\mu$) à valeurs dans un espace X est *faiblement analytique* si quelle que soit la fonctionnelle linéaire $f \in X^*$, la fonction $f$ ($x$ ($\mu$)) est analytique dans l'acception ordinaire de ce terme.

Un opérateur $P$ est *analytique* s'il transforme toute fonction faiblement analytique $x$ ($\mu$) en une fonction faiblement analytique $P$ ($x$ ($\mu$)) (dans un espace Y).

Dans l'équation (19), supposons que $\mu$ est un paramètre numérique, $\pi$ et $R$, des opérateurs analytiques. Alors, on vérifie immédiatement que toutes les approximations (21) sont des fonctions analytiques de $\mu$. Soit $f$ une fonctionnelle linéaire continue dans l'espace X. Les normes des éléments $\tilde{x}_n$ ($\mu$) étant bornées uniformément (en $\mu$), les fonctions analytiques

$$\varphi_n(\mu) = f(\tilde{x}_n(\mu)) \quad (n = 0, 1, \ldots)$$

sont bornées dans leur ensemble, donc la fonction

$$\varphi(\mu) = \lim_{n\to\infty} \varphi_n(\mu) = f(x^*(\mu))$$

est analytique. Comme $f$ est arbitraire, $x^*$ ($\mu$) est une fonction faiblement analytique de $\mu$.

Si l'on définit de façon naturelle le développement de la fonction $x^*$ ($\mu$) en série suivant les puissances de $\mu$

$$x^*(\mu) = x_0 + \mu x_1 + \ldots + \mu^n x_n + \ldots,$$

de ce qui vient d'être dit il résulte que cette série est faiblement convergente pour tous les $|\mu| < \mu_0$.

Un grand nombre de travaux ont été consacrés à la méthode de Newton, ses applications et généralisations. Citons en particulier Krasnosselski et autres, Collatz et les références indiquées dans ces ouvrages.

## § 3. Application de la méthode de Newton à des équations fonctionnelles concrètes

**3.1.** Commençons par l'application de la méthode de Newton à une équation réelle ou complexe

$$f(z) = 0. \tag{1}$$

Le théorème 1.6 dans lequel•

$$\eta \geqslant \frac{|f(z_0)|}{|f'(z_0)|}, \quad K \geqslant \max \frac{|f''(z)|}{|f'(z_0)|},$$

où $z_0$ est l'approximation initiale, nous dit que cette équation admet une racine si

$$h=\eta K\leqslant\frac{1}{2}\quad\text{ou}\quad\frac{|f(z_0)|\,|f''(z_0)|}{|f'(z_0)|^2}\leqslant\frac{1}{2},$$

cette racine est située dans le disque

$$|z-z_0|\leqslant r_0=\frac{1-\sqrt{1-2h}}{h}\eta \tag{2}$$

et est unique dans le disque

$$|z-z_0|<r_1=\frac{1+\sqrt{1-2h}}{h}\eta\,^{*}). \tag{3}$$

La méthode de Newton peut être utilisée pour résoudre des systèmes d'équations algébriques

$$\varphi_j(\xi_1, \xi_2, \ldots, \xi_m)=0 \quad (j=1, 2, \ldots, m), \tag{4}$$

que nous traiterons comme une équation

$$P(x)=0$$

dans un espace X de dimension $m$:

$$y=P(x)$$

$(y=(\varphi_1(\xi_1, \ldots, \xi_m), \ldots, \varphi_m(\xi_1, \ldots, \xi_m)), \; x=(\xi_1, \ldots, \xi_m))$.

Soit $x_0=(\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_m^{(0)})$ une approximation initiale. En la portant dans l'équation de la méthode de Newton

$$P'(x_0)(x-x_0)+P(x_0)=0$$

et en tenant compte de l'expression de la dérivée $P'(x_0)$ indiquée dans XVII.1.6, on obtient pour l'écart $\Delta x=x_1-x_0=$ $=(\Delta\xi_1, \Delta\xi_2, \ldots, \Delta\xi_m)$ le système d'équations linéaires

$$\sum_{k=1}^{m}\frac{\partial\varphi_j(\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_m^{(0)})}{\partial\xi_k}\Delta\xi_k=-\varphi_j(\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_m^{(0)}) \tag{5}$$

$$(j=1, 2, \ldots, m),$$

d'où l'on déduit $\Delta x$, donc $x_1$. On trouve de même $x_2$, etc.

Si l'on se sert du processus modifié, la matrice du système (5) reste inchangée à chaque itération, seuls changent les seconds membres.

Les conditions de convergence du processus dépendent de la façon dont est introduite la norme dans l'espace X. Nous traitons deux cas: $\mathrm{X}=\mathrm{l}_m^{\infty}$ et brièvement $\mathrm{X}=\mathrm{l}_m^{2}$.

*) Ceci étant, on suppose que la dérivée seconde est majorable dans le disque $|z-z_0|\leqslant r$, où $r\geqslant r_0$ si l'on envisage (2), et $r\geqslant r_1$ si l'on envisage (3).

THÉORÈME 1. *Supposons que les fonctions $\varphi_1, \varphi_2, \ldots, \varphi_m$ et l'approximation initiale $x_0$ vérifient les conditions:*

1) $|\varphi_j(\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_m^{(0)})| \leqslant \eta' \quad (j=1, 2, \ldots, m)$;

2) *le déterminant $D$ de la matrice de Jacobi*

$$\left(\frac{\partial\varphi_j}{\partial\xi_k}(\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_m^{(0)})\right) \quad (j, k=1, 2, \ldots, m)$$

*n'est pas nul et si $A_{j,k}$ est le cofacteur de $\frac{\partial\varphi_j}{\partial\xi_k}$, alors*

$$\frac{1}{|D|}\sum_{j=1}^{m} |A_{j,k}| \leqslant B' \quad (k=1, 2, \ldots, m);$$

3) $\left|\frac{\partial^2\varphi_j}{\partial\xi_k\,\partial\xi_s}(\xi_1, \xi_2, \ldots, \xi_m)\right| \leqslant L \; (j, k, s=1, 2, \ldots, m; \; |\xi_i - \xi_i^{(0)}| \leqslant r)$;

4) $h = B'^2\eta' L m^2 \leqslant \frac{1}{2}$.

*Si*

$$r \geqslant \frac{1-\sqrt{1-2h}}{h} B'h', \tag{6}$$

*alors le système* (4) *admet une solution $x^* = (\xi_1^*, \xi_2^*, \ldots, \xi_m^*)$ située au voisinage de $x_0$, vers laquelle tendent les processus principal et modifié de Newton.*

Prouver ce théorème revient à vérifier les conditions du théorème 1.6, plus exactement. la remarque 1 qui suit ce théorème, si $X = l_m^\infty$. Cette vérification est laissée au soin du lecteur *).

REMARQUE 1. Les vitesses de convergence du processus principal et du processus modifié admettent les majorations (33) et (34) du théorème 1.6.

REMARQUE 2. Comme pour $0 < h \leqslant \frac{1}{2}$

$$\frac{1-\sqrt{1-2h}}{h} \leqslant 2,$$

(6) a lieu si $r = 2B'\eta'$.

REMARQUE 3. Soit $m=2$ le nombre d'équations du système (4). On a

$$A_{j,k} = \pm\frac{\partial\varphi_{3-k}}{\partial\xi_{3-j}}(\xi_1^{(0)}, \xi_2^{(0)}) \quad (j, k=1, 2),$$

donc, si l'on connaît la majoration

$$\left|\frac{\partial\varphi_j}{\partial\xi_k}(\xi_1^{(0)}, \xi_2^{(0)})\right| \leqslant l \quad (j, k=1, 2),$$

*) Il est immédiat que $\| P''(x) \| \leqslant Lm^2$ (cf. XVII.2.2).

on peut prendre pour $B'$

$$B' = \frac{2l}{|D|},$$

et la condition 4) devient

$$\frac{16l^2\eta' L}{|D|} \leqslant \frac{1}{2}$$

ou

$$32l^2\eta' L \leqslant |D|.$$

Cette condition a été établie par A. Ostrovski. Il est curieux de noter que la démonstration directe de la convergence de la méthode de Newton est très compliquée même dans ce cas élémentaire.

Si l'on utilise la normalisation de l'espace $l_m^2$, il faut remplacer les conditions 1) à 4) du théorème 1 par

1') $\|\Gamma_0(P(x_0))\| = \|\Delta x\| = [\sum_{k=1}^{m} |\Delta\xi_k|^2]^{1/2} \leqslant \eta$;

2') $\|\Gamma_0\| = \sqrt{\Lambda} \leqslant B'$;

où $\Gamma_0$ représente comme toujours $[P'(x_0)]^{-1}$ et $\Lambda$ la plus grande valeur propre de la matrice $\Gamma_0\Gamma_0^*$; pour $B'$ on peut prendre

$$B' = \frac{1}{|D|}\left[\sum_{j=1}^{m}\sum_{k=1}^{m} |A_{j,k}|^2\right]^{1/2};$$

3') $\|P''(x)\| \leqslant \left[\sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{s=1}^{m}\left|\frac{\partial^2\varphi_j}{\partial\xi_k\,\partial\xi_s}\right|^2\right]^{1/2} \leqslant K'$;

pour $K'$ on peut donc prendre

$$K' = Lm\sqrt{m};$$

4') $h = B'K'\eta \leqslant \frac{1}{2}$.

La réalisation de ces conditions assure l'existence de la solution du système (4) et la convergence vers elle des processus principal et modifié de Newton.

Illustrons ceci sur un exemple. Soit le système

$$3\xi_1^2\xi_2 + \xi_2^3 = 1,$$
$$\xi_1^4 + \xi_1\xi_2^3 = 1.$$

Prenons pour approximation initiale $x_0 = (\xi_1^{(0)}, \xi_2^{(0)})$:

$$\xi_1^{(0)} = 0{,}98, \quad \xi_2^{(0)} = 0{,}32.$$

Pour la détermination des écarts on obtient le système

$$1{,}880\Delta\xi_1 + 3{,}188\Delta\xi_2 = 0{,}045,$$
$$3{,}798\Delta\xi_1 + 0{,}301\Delta\xi_2 = 0{,}046,$$

d'où l'on trouve

$$\Delta\xi_1 = 0{,}0105, \quad \Delta\xi_2 = 0{,}0075.$$

Il s'ensuit

$$-\Gamma_0 = \begin{pmatrix} -0{,}026 & 0{,}256 \\ 0{,}329 & -0{,}163 \end{pmatrix}.$$

Calculons les constantes $B'$, $\eta'$ et $L$ du théorème 1 en utilisant la normalisation de l'espace $l_2^\infty$. On trouve

$$\| \Gamma_0 \| = 0{,}492 < 0{,}5 = B',$$
$$\| P(x_0) \| = \max [0{,}046;\ 0{,}045] < 0{,}05 = \eta'.$$

Majorons les dérivées des fonctions $\varphi_1$ et $\varphi_2$ pour $\| x - x_0 \| < 2B'\eta'$, c'est-à-dire pour $0{,}93 \leqslant \xi_1 \leqslant 1{,}03$; $0{,}27 \leqslant \xi_2 \leqslant 0{,}37$. Comme

$$\frac{\partial^2\varphi_1}{\partial\xi_1^2} = 6\xi_2, \quad \frac{\partial^2\varphi_1}{\partial\xi_1\,\partial\xi_2} = 6\xi_1, \quad \frac{\partial^2\varphi_1}{\partial\xi_2^2} = 6\xi_2,$$
$$\frac{\partial^2\varphi_2}{\partial\xi_1^2} = 12\xi_1^2, \quad \frac{\partial^2\varphi_2}{\partial\xi_1\,\partial\xi_2} = 3\xi_2^2, \quad \frac{\partial^2\varphi_2}{\partial\xi_2^2} = 6\xi_1\xi_2,$$

pour $L$ on peut prendre

$$L = 12{,}8 > \max \frac{\partial^2\varphi_2}{\partial\xi_1^2} = 12\,(1{,}03)^2.$$

Alors

$$h = B'^2\eta' L m^2 = 0{,}25\cdot 0{,}05\cdot 12{,}8\cdot 4 = 0{,}64.$$

Donc, le théorème 1 ne permet pas d'établir la convergence du processus de Newton. Ceci est dû à la trop grossière majoration de $\| P''(x) \|$. En effet, nous avons pris

$$\| P''(x) \| \leqslant Lm^2 = 51{,}2,$$

alors que

$$\| P''(x) \| \leqslant 16 = K'.$$

Ceci nous donne

$$h = B'^2\eta' K' = 0{,}25\cdot 0{,}05\cdot 16 = 0{,}2,$$

donc le processus de Newton est convergent.

A noter que si l'on se sert directement du théorème 1.6, on peut obtenir une valeur plus petite encore pour $h$. En effet,

$$\| \Gamma_0 (P(x_0)) \| = \| \Delta x \| = \max [|\Delta\xi_1|,\ |\Delta\xi_2|] = 0{,}0105$$

et

$$\| \Gamma_0 P''(x) \| < 7{,}2.$$

Donc, si $\eta = 0{,}0105$, $K = 7{,}2$, on obtient

$$h = \eta K < 0{,}08. \qquad (7)$$

S'agissant du processus modifié, on obtient les approximations successives

$$\xi_1^{(1)} = 0{,}9905, \quad \xi_2^{(1)} = 0{,}3275,$$
$$\xi_1^{(2)} = 0{,}99117, \quad \xi_2^{(2)} = 0{,}32738,$$
$$\xi_1^{(3)} = 0{,}991189, \quad \xi_2^{(3)} = 0{,}327382.$$

La majoration de l'erreur de la troisième approximation par la formule (34) du théorème 1.6 est égale à 0,0008 (pour $h$ on a pris la valeur (7)).

En se servant du théorème 2.4, on peut exhiber une majoration plus exacte encore de l'erreur, notamment

$$0{,}991173 \leqslant \xi_1^* \leqslant 0{,}991205, \quad 0{,}327366 \leqslant \xi_2^* \leqslant 0{,}327398.$$

Donc, en troisième approximation, nous avons quatre décimales vraies.

Si l'on fait intervenir la norme de l'espace $l_2^2$, on obtient à partir des conditions 1') à 3')

$$\| \Gamma_0 \| \leqslant 0{,}448 < 0{,}5 = B',$$
$$\| \Gamma_0 (P(x_0)) \| = \| \Delta x \| = 0{,}0133 < 0{,}015 = \eta,$$
$$\| P''(x) \| < 15{,}2 = K', \quad h = B'K'\eta = 0{,}1026.$$

3.2. Appliquons la méthode de Newton aux équations intégrales non linéaires. Soit donnée l'équation intégrale

$$x(s) = \int_0^1 K(s,\ t,\ x(t))\, dt, \tag{8}$$

où $K(s, t, u)$ est une fonction continue en tous ses arguments, admettant autant de dérivées continues que l'on veut. Soit dans l'espace fonctionnel X l'opérateur $P$:

$$y = P(x), \quad y(s) = x(s) - \int_0^1 K(s,\ t,\ x(t))\, dt. \tag{9}$$

Mettons l'équation (8) sous la forme de l'équation fonctionnelle

$$P(x) = 0.$$

Le processus de Newton pour cette équation se construit de la manière suivante: soit $x_0$ l'approximation initiale; supposons que l'espace X et la fonction $K(s, t, u)$ sont tels que $P'(x_0)$ puisse être obtenue par « dérivation sous le signe somme», c'est-à-dire $z = P'(x_0)(x)$ signifie que

$$z(s) = x(s) - \int_0^1 K'_u(s,\ t,\ x_0(t))\, x(t)\, dt\ ; \tag{10}$$

l'erreur $\Delta x = x_1 - x_0$ se déduit à partir de l'équation

$$P'(x_0)(\Delta x) = -P(x_0)$$

qui, en vertu de (10), s'écrit

$$\Delta x(s) - \int_0^1 K'_u(s,\ t,\ x_0(t))\, \Delta x(t)\, dt = \varepsilon_0(s), \tag{11}$$

où

$$\varepsilon_0(s) = \int_0^1 K(s,\ t,\ x_0(t))\, dt - x_0(s)$$

est le résidu de l'équation (8), correspondant à l'approximation initiale.

Donc, pour trouver l'approximation suivante, il faut résoudre une équation intégrale linéaire ; à noter que dans le processus modifié on aura toujours le même noyau.

La convergence du processus de Newton dépend de l'espace dans lequel on se place. Ainsi, dans $X = C\,[0,\ 1]$ on a le

THÉORÈME 2. *Soient réalisées les conditions:*

1) *pour le noyau* $K(s,\ t) = K'_u(s,\ t,\ x_0(t))$, *l'équation intégrale* (11) *admet la résolvante* $G(s,\ t)$ *et de plus*

$$\int_0^1 |G(s,\ t)|\,dt \leqslant$$ [redacted]

2) $|\varepsilon_0(s)| \leqslant \eta'$ ( [redacted]

3) $\int_0^1 |K''_{u^2}(s,\ t,\ u)|\,dt \leqslant K'$

*dans les domaines de valeurs de s et u, définis par*

$$0 \leqslant s \leqslant 1, \quad |u - x_0(s)| \leqslant 2\,(1 + B)\,\eta';$$

4) $h = (1+B)^2 K'\eta' \leqslant \frac{1}{2}$.

*Alors le processus de Newton (principal et modifié) converge vers la solution de l'équation* (8). *Ceci étant, la solution* $x^*(s)$ *est située dans le domaine*

$$|x^*(s) - x_0(s)| \leqslant \frac{1-\sqrt{1-2h}}{h}(1+B)\,\eta' \quad (s \in [0,\ 1])$$

*et est unique dans le domaine*

$$0 \leqslant s \leqslant 1, \quad |u - x_0(s)| \leqslant \frac{1+\sqrt{1-2h}}{h}(1+B)\,\eta'.$$

On vérifie sans peine que les conditions du théorème 1.6 son réalisées si l'on tient compte du fait que d'après le théorème XVII.3. la dérivée $P'(x_0)$ de l'opérateur $P$ est de la forme (10) et $P''(x)$ est u opérateur intégral bilinéaire de noyau $K'_{u^2}(s,\ t,\ x_0(t))$. L'opérateu $\Gamma_0 = [P'(x_0)]^{-1}$, comme indiqué dans XIII.6.1, est de la forme

$$z = \Gamma_0(y), \quad z(s) = y(s) + \int_0^1 G(s,\ t)\,y(t)\,dt,$$

donc

$$\|\Gamma_0\| \leqslant 1 + B.$$

Si $X = L^2(0, 1)$, en utilisant le théorème XVII.3.1, on peut indiquer d'autres conditions de solubilité de l'équation intégrale (8). Plus exactement

1') $\int_0^1 |\varepsilon_0(s)|^2 ds = \eta'^2$;

2') $\left|\frac{\lambda_n}{1-\lambda_n}\right| \leqslant B'$ $(n = 1, 2, \ldots)$,

où $\lambda_n$ sont les valeurs caractéristiques du noyau $K(s, t) = K'_u(s, t, x_0(t))$ s'il est symétrique (s'il n'était pas symétrique, cette condition se compliquerait; cf. IV.2.6 et V.2.7 du tome 1);

3') $|K''_{u^2}(s, t, u)| \leqslant K'$ $(s, t \in [0, 1], -\infty < u < \infty)$;

4') $h = B'^2 K' \eta' \leqslant \frac{1}{2}$.

**3.3.** Les conditions du théorème 2 ou les conditions 1') à 4') supposent une approximation initiale $x_0$ assez proche de la solution exacte. Nous allons voir maintenant une méthode de construction assez générale d'une telle approximation initiale, consistant à remplacer l'équation intégrale (8) par une équation de structure plus simple dont la résolution se ramène à celle d'un système algébrique.

On suppose notamment que le noyau $K(s, t, u)$ peut être approché par un noyau de forme plus simple

$$H(s, t, u) = \sum_{k=1}^{m} h_k(t, u)\,\omega_k(s),$$

où $\omega_k(s)$ $(k = 1, 2, \ldots, m)$ sont des fonctions que, sans nuire à la généralité, on peut supposer être deux à deux orthogonales et normées. Pour $H$ on peut prendre, par exemple, une série partielle de Fourier de la fonction $K$ suivant le système orthonormal complet $\{\omega_k(s)\}$ $(k = 1, 2, \ldots)$.

L'équation

$$x(s) = \int_0^1 H(s, t, x(t))\,dt \tag{13}$$

doit naturellement être considérée comme proche de l'équation (8). Or, la solution $x_0(t)$ de l'équation (13) s'écrit

$$x_0(t) = \sum_{k=1}^{m} A_k \omega_k(t),$$

où les coefficients $A_k$ $(k = 1, 2, \ldots, m)$ sont déterminés à partir du système algébrique non linéaire

$$A_j = \int_0^1 h_j\Big(t, \sum_{k=1}^{m} A_k \omega_k(t)\Big)\, dt \quad (j = 1, 2, \ldots, m). \tag{14}$$

C'est la fonction $x_0(t)$ que l'on prendra pour approximation initiale de l'équation (8).

On se servira du schéma du théorème 2.6, car dans de nombreux cas, seuls l'existence et le domaine de disposition et d'unicité de la solution de l'équation (8) présentent de l'intérêt. On posera donc $\mathbf{X} = \mathbf{C}\,[0, 1]$ et l'on mettra l'équation (8) sous la forme

$$P(x) \equiv \pi(x) + \mu R(x) = 0,$$

$$\pi(x)(s) = x(s) - \int_0^1 H(s, t, x(t))\, dt,$$

$$R(x)(s) = \int_0^1 [K(s, t, x(t)) - H(s, t, x(t))]\, dt,$$

$$\mu = 1.$$

On a alors le théorème suivant.

THÉORÈME 3. *Soient réalisées les conditions:*

1) *la résolvante* $G_0(s, t)$ *du noyau* $H'_u(s, t, x_0(t))$ *existe et*

$$\int_0^1 |G_0(s, t)|\, dt \leqslant B \quad (s \in [0, 1]);$$

2) $$\left| \int_0^1 [K(s, t, x_0(t)) - H(s, t, x_0(t))]\, dt \right| = |\varepsilon_0(s)| \leqslant \eta \quad (s \in [0, 1]);$$

3) $$\int_0^1 |K'_u(s, t, x_0(t)) - H'_u(s, t, x_0(t))|\, dt \leqslant \alpha \quad (s \in [0, 1]);$$

4) $$\int_0^1 |H''_{u^2}(s, t, u)\, dt \leqslant K \quad (s \in [0, 1],\ |u - x_0(s)| \leqslant r);$$

5) $$\int_0^1 |K''_{u^2}(s, t, u) - H''_{u^2}(s, t, u)|\, dt \leqslant L \quad (s \in [0, 1],\ |u - x_0(s)| \leqslant r);$$

6) $$h = \frac{(1+B)^2 (K+L)\, \eta}{[1 - \alpha(B+1)]^2} \leqslant \frac{1}{2}, \quad \alpha(B+1) < 1.$$

*Si*

$$r \geqslant r_0 = \frac{1-\sqrt{1-2h}}{h} \frac{\eta(1+B)}{1-\alpha(1+B)},$$

*alors l'équation* (8) *admet la solution* $x^*$ *et de plus*

$$|x^*(s) - x_0(s)| \leqslant r_0. \tag{15}$$

L'avantage du théorème énoncé sur le théorème 2 est que le noyau $H'_u(s, t, x_0(t))$ est dégénéré, donc la résolvante $G_0(s, t)$ est facile à trouver.

Pour trouver la solution de l'équation (8) on peut utiliser le processus indiqué dans 2.6, bien que celui-ci ne converge pas aussi vite que le processus modifié.

Illustrons tout ceci sur un exemple. Soit donnée l'équation

$$x(s) = \frac{1}{2}\int_0^1 [x(t)]^2 \sin st\, dt + 1.$$

Pour noyau approximant on prend

$$H(s, t, u) = \frac{1}{2} stu^2 + 1,$$

l'équation (13) devient alors

$$x(s) = \frac{1}{2}\int_0^1 [x(t)]^2 st\, dt + 1.$$

La solution de cette équation est $x_0(t) = 1 + At$, $A$ étant racine d'une équation du second degré, soit $A = 0{,}405887$. On calcule sans peine les constantes du théorème 3:

$$B = 1{,}124; \quad \eta = 0{,}0367; \quad \alpha = 0{,}055; \quad K = 0{,}5; \quad L = 0{,}0417,$$

d'où

$$0{,}1 < h < 0{,}103; \quad \alpha(1+B) < 0{,}12$$

et d'après (15)

$$|x^*(s) - x_0(s)| < 0{,}096 \quad (s \in [0, 1]).$$

Si on avait pris

$$H(s, t, u) = \frac{1}{2}\left(st - \frac{s^3t^3}{6}\right) u^2 + 1,$$

on aurait obtenu

$$x_0(s) = 1 + 0{,}38617s - 0{,}0345s^3,$$

et la majoration (15) aurait été de la forme

$$|x^*(s) - x_0(s)| < 0{,}0119 \quad (s \in [0, 1]).$$

3.4. Etudions le problème de l'application de la méthode de Newton aux équations différentielles. Soit donnée l'équation différentielle

$$x'(t) - \varphi(x(t), t) = 0 \quad (x(0) = 0). \tag{16}$$

Munissons l'espace $C^1$ des fonctions continûment dérivables sur l'intervalle $[0, a]$, nulles pour $t = 0$, de la norme

$$\| x \| = \max_{t \in [0, a]} | x(t) | + \lambda \max_{t \in [0, a]} | x'(t) |,$$

où $\lambda$ est un coefficient positif qui sera déterminé par la suite. Supposons que la fonction $\varphi(u, t)$ est continue et possède une dérivée seconde par rapport à $u$, continue dans le domaine

$$0 \leqslant t \leqslant a, \quad | u - x_0(t) | \leqslant \delta \quad (x_0 \in C^1),$$

et considérons l'opérateur $P$

$$y = P(x), \quad y(t) = x'(t) - \varphi(x(t), t).$$

Il est immédiat de vérifier que $P$ est un opérateur de la boule $\| x - x_0 \| \leqslant \delta$ (que nous désignons par $\Omega$) dans l'espace $C[0, 1]$ et possède dans $\Omega$ des dérivées première et seconde continues, telles que

$$P'(z)(x)(t) = x'(t) - \varphi'_u(z(t), t)\, x(t), \tag{17}$$

$$P''(z)(x, \tilde{x})(t) = -\varphi''_{u^2}(z(t), t)\, x(t)\, \tilde{x}(t). \tag{18}$$

Trouvons $\Gamma_0 = [P'(x_0)]^{-1}$. D'après (17), l'élément $x = \Gamma_0(y)$ est solution de l'équation différentielle

$$x'(t) = \varphi'_u(x_0(t), t)\, x(t) + y(t), \tag{19}$$

c'est-à-dire

$$x(t) = \psi(t) \int_0^t \frac{y(s)}{\psi(s)}\, ds,$$

où

$$\psi(t) = \exp\left( \int_0^t \varphi'_u(x_0(s), s)\, ds \right).$$

Il s'ensuit

$$\max_{t \in [0, a]} | x(t) | \leqslant a \frac{\max | \psi(t) |}{\min | \psi(t) |} \| y \|.$$

De l'équation (19) il vient

$$\max_{t \in [0, a]} [x'(t) | \leqslant \max_{t \in [0, a]} | \varphi'_u(x_0(t), t) | \max_{t \in [0, a]} | x(t) | + \| y \| \leqslant \theta \| y \|,$$

donc

$$\| x \| \leqslant \left[ a \frac{\max | \psi(t) |}{\min | \psi(t) |} + \lambda\theta \right] \| y \|,$$

et

$$\| \Gamma_0 \| \leqslant a \frac{\max | \psi(t) |}{\min | \psi(t) |} + \lambda\theta.$$

Cette majoration nous permet d'appliquer le théorème 1.6 qui nous conduit au

THÉORÈME 4. *Soient réalisées les conditions*:

1) $|x_0'(t) - \varphi(x_0(t), t)| \leqslant \eta'$ $(t \in [0, a])$;

2) $|\varphi_u'(x_0(t), t)| \leqslant M_1$ $(t \in [0, a])$;

3) $|\varphi_{u^2}''(u, t)| \leqslant M_2$ $(t \in [0, a], |u - x_0(t)| \leqslant \delta)$;

4) $h_0 = M_2 a^2 e^{4aM_1} \eta' < \frac{1}{2}$.

*Si*

$$\delta > r_0 = \frac{1 - \sqrt{1 - 2h_0}}{h_0} e^{2aM_1} a\eta',$$

*alors l'équation* (16) *possède dans l'intervalle* $[0, a]$ *une solution* $x^*(t)$ *unique telle que*

$$|x^*(t) - x_0(t)| < r_0.$$

Pour vérifier les conditions du théorème 1.6 il suffit de s'assurer que

$$\max |\psi(t)| \leqslant e^{aM_1}, \quad \min |\psi(t)| \geqslant e^{-aM_1},$$

donc, pour $B'$ on peut prendre

$$B' = ae^{2aM_1} + \lambda\theta.$$

On peut ensuite poser $K' = M_2$. Donc, pour que l'équation (16) admette une solution il faut que

$$h = B'^2 K' \eta' = (ae^{2aM_1} + \lambda\theta)^2 M_2 \eta' \leqslant \frac{1}{2}. \tag{20}$$

Pour $\lambda$ assez petit, cette inégalité découle de la condition 4), puisque $\lim\limits_{\lambda \to 0} h = h_0$.

REMARQUE 1. Le cas d'une condition initiale différente de zéro se ramène au cas étudié par un changement de variables évident.

REMARQUE 2. La méthode exposée est valable pour l'étude de systèmes d'équations différentielles. Les difficultés de son application au cas étudié tiennent à l'impossibilité, en général, de trouver la solution analytique du système linéaire, donc de majorer la norme de l'opérateur $\Gamma_0$.

3.5. Considérons maintenant l'équation différentielle

$$x''(t) + x(t) + \mu\varphi(x(t), x'(t), t) = 0, \tag{21}$$

où $\varphi(u, v, t)$ est une fonction continue avec ses dérivées première et seconde par rapport à $u$ et $v$, périodique en $t$, de période $\omega > 0$.

En procédant comme dans 3.4, indiquons les conditions sous lesquelles l'équation (21) admet une solution $\omega$-périodique, c'est-à-dire

vérifiant les conditions aux limites

$$x(\omega) = x(0), \quad x'(\omega) = x'(0). \tag{22}$$

Nous utilisons cette fois le schéma du théorème 2.6. Considérons l'espace $\widetilde{C}^2$ des fonctions bicontinûment dérivables périodiques (de période $\omega$). Munissons $\widetilde{C}^2$ de la norme

$$\| x \| = \max_t |x(t)| + \max_t |x'(t)| + \lambda \max_t |x''(t)| \quad (x \in \widetilde{C}^2),$$

où $\lambda > 0$ sera déterminé par la suite. Soient par ailleurs les opérateurs $\pi$ et $R$ :

$$y = \pi(x), \quad y(t) = x''(t) + x(t), \quad z = R(x), \quad z(t) = \varphi(x(t), x'(t), t).$$

L'équation (21) s'écrit alors

$$\pi(x) + \mu R(x) = 0. \tag{23}$$

Effectuons les majorations nécessaires pour appliquer le théorème 2.6.

L'opérateur $\pi$ est manifestement continu et linéaire, donc la solution $x_0$ de l'équation $\pi(x) = 0$ est l'élément $x_0 = 0$. Si $\cos \omega \neq 1$, il existe $\Gamma_0 = \pi^{-1}$. On détermine l'opérateur $\Gamma_0$ en résolvant l'équation différentielle

$$x'' + x = y$$

avec les conditions (22). Nous omettons les calculs peu compliqués mais laborieux et nous indiquons directement la majoration de la norme de l'opérateur $\Gamma_0$:

$$\| \Gamma_0 \| \leqslant \omega \left[ 2 + \frac{1 + |\cos \omega| + |\sin \omega|}{1 - \cos \omega} \right] + \lambda\theta \leqslant$$
$$\leqslant \frac{4 + 1 + \sqrt{2}}{1 - \cos \omega}\, \omega + \lambda\theta < \frac{6,5\, \omega}{1 - \cos \omega} + \lambda\theta,$$

où $\theta$ est une expression contenant $\omega$.

Donc, pour $B$ on peut prendre

$$B = \frac{6,5\, \omega}{1 - \cos \omega} + \lambda\theta = B_0 + \lambda\theta.$$

Comme

$$R'(x_0)(x)(t) = \varphi'_u(0, 0, t)\, x(t) + \varphi'_v(0, 0, t)\, x'(t),$$
$$R''(z)(x, \widetilde{x})(t) = \varphi''_{u^2}(z(t), z'(t), t)\, x(t)\, \widetilde{x}(t) +$$
$$+ \varphi''_{uv}(z(t), z'(t), t)\, [x(t)\, \widetilde{x}'(t) + x'(t)\, \widetilde{x}(t)] +$$
$$+ \varphi''_{v^2}(z(t), z'(t), t)\, x'(t)\, \widetilde{x}'(t),$$

on obtient le

THÉORÈME 5. *Soient réalisées les conditions:*

1) $\omega = 2n\pi$ $(n = 1, 2, \ldots)$;

2) $|\varphi(0, 0, t)| \leqslant \eta$;

3) $|\varphi'_u(0, 0, t)| \leqslant \alpha$, $|\varphi'_v(0, 0, t)| \leqslant \alpha$;

4) $|\varphi''_{u^2}(u, v, t)| \leqslant L$, $|\varphi''_{uv}(u, v, t)| \leqslant L$, $|\varphi''_{v^2}(u, v, t)| \leqslant L$.

*Si*

$$\mu < \frac{1-\cos\omega}{6{,}5\,\omega(\alpha + \sqrt{2L\eta})}, \tag{24}$$

*alors l'équation* (21) *possède une solution unique périodique, de période* $\omega$.

Pour prouver ce théorème, il suffit de vérifier que l'inégalité (20) du théorème 2.6, qui s'écrit ici

$$h_\mu = \frac{B^2 L\eta\mu^2}{(1-\alpha B\mu)^2} \leqslant \frac{1}{2},$$

est réalisée si $\lambda$ est assez petit et

$$\frac{B_0^2 L\eta\mu^2}{(1-\alpha B_0\mu)^2} < \frac{1}{2}.$$

Or. cette inégalité équivaut à (24).

REMARQUE. Si la fonction $\varphi$ est analytique en ses arguments, d'après ce qui a été dit dans 2.7, on peut conclure que la solution périodique de l'équation (21) sera analytique en $\mu$ pour les $\mu$ vérifiant (24).

Ces raisonnements sont valables pour des équations d'ordre supérieur.

3.6. Appliquons la méthode de Newton à la théorie des perturbations.

Soient $U$ et $V$ des opérateurs linéaires continus d'un $B$-espace $\mathbf{X}$ dans lui-même. On demande une valeur propre et l'élément propre associé de l'opérateur $U_t = U + tV$, en supposant que la solution de ce problème est connue pour $t = 0$, c'est-à-dire est connue la valeur propre $\lambda_0$ de $U$ et l'élément propre $x_0$ qui lui est associé.

Nous allons imposer des conditions accessoires aux opérateurs $U$ et $V$ et prouver que pour des $t$ assez petits, l'opérateur $U_t$ possède une valeur propre $\lambda_t$ et l'élément propre $x_t$, proches respectivement de $\lambda_0$ et $x_0$. Plus exactement, on a le

THÉORÈME 6. *Soient réalisées les conditions:*

1) $\overline{\lambda}_0$ *est valeur propre de l'opérateur* $U^*$:

$$U^*(f_0) = \overline{\lambda}_0 f_0,$$

$f_0$ *désignant un élément propre normé par la condition*

$$f_0(x_0) = 1;$$

2) *l'opérateur* $U - \lambda_0 I$ *dont le domaine de valeurs est situé dans l'espace* $X_0 = N(f_0)$ *des* $y \in X$ *tels que* $f_0(y) = 0$ (cf. XII.2.1), *traité comme un opérateur de* $X_0$ *dans* X, *admet un inverse T continu*

$$T(Ux - \lambda_0 x) = x, \quad UTx - \lambda_0 Tx = x \quad (x \in X_0).$$

*Si*

$$\|T\|\,\|V\|\,|t| \leqslant \frac{1}{1+2\|f_0\|\,\|x_0\|+2\sqrt{(1+\|f_0\|\,\|x_0\|)\,\|f_0\|\,\|x_0\|}}, \tag{25}$$

*alors l'opérateur* $U_t$ *possède une valeur propre* $\lambda_t$ *telle que, si l'élément propre associé* $x_t$ *est normé par la condition*

$$f_0(x_t) = 1,$$

*alors*

$$|\lambda_t - \lambda_0| \leqslant A\,|t|, \quad \|x_t - x_0\| \leqslant B\,|t|,$$

*où A et B sont des constantes déterminées par les opérateurs U et V.*

Démonstration. On demande de déterminer $\lambda$ et $x$ à partir des conditions

$$U(x) + tV(x) = \lambda x, \quad f_0(x) = 1.$$

Faisons le changement

$$x = x_0 + z, \quad \lambda = \lambda_0 + \delta.$$

Les nouvelles variables doivent vérifier les conditions

$$U(z) + tV(x_0 + z) = \delta(x_0 + z) + \lambda_0 z, \quad f_0(z) = 0,$$

c'est-à-dire on demande de trouver dans l'espace $X_0$ la solution de l'équation

$$U(z) + tV(x_0 + z) = \delta(x_0 + z) + \lambda_0 z. \tag{26}$$

Appliquons la fonctionnelle $f_0$ aux deux membres de l'équation (26). Comme

$$f_0(Uz - \lambda_0 z) = U^*(f_0)(z) - (\overline{\lambda}_0 f_0)(z) = 0,$$

on est conduit à la relation

$$\delta = tf_0(V(x_0 + z)), \tag{27}$$

exprimant $\delta$ en fonction de $z$. En portant (27) dans (26), on obtient pour la détermination de $z$ l'équation

$$U(z) - \lambda_0 z + t\,[V(x_0 + z) - f_0(V(x_0 + z))(x_0 + z)] = 0. \tag{28}$$

Les deux membres de cette équation appartiennent au sous-espace $X_0$, on peut donc leur appliquer l'opérateur $T$; on obtient

$$\pi(z) + tR(z) = 0, \tag{29}$$

où

$$\pi(z) = z,$$
$$R(z) = T\,[V(x_0 + z) - f_0(V(x_0 + z))(x_0 + z)].$$

En appliquant le théorème 2.6 à l'équation (29), on obtient ($z_0 = 0$)

$$\pi'(z_0) = I, \quad \Gamma_0 = I, \quad \pi''(z) = 0,$$
$$R(z_0) = T\,[V(x_0) - f_0(V(x_0))\,x_0],$$
$$R'(z)(x) = T\,[V(x) - f_0(V(x))(x_0 + z) - f_0(V(x_0 + z))\,x],$$
$$R'(z)(x, \tilde{x}) = -T\,[f_0(V(x))\,\tilde{x} + f_0(V(x))\,\tilde{x}] =$$
$$= -f_0(V(x))\,T(\tilde{x}) - f_0(V(\tilde{x}))\,T(x).$$

On peut donc prendre

$$\eta = \| T \| \, \| V \| \, [1 + \| f_0 \| \, \| x_0 \|] \, \| x_0 \|, \quad B = 1,$$
$$\alpha = \| T \| \, \| V \| \, [1 + 2 \, \| f_0 \| \, \| x_0 \|],$$
$$K = 0, \quad L = 2 \, \| T \| \, \| V \| \, \| f_0 \|.$$

La condition de solubilité de l'équation (29) s'écrit alors

$$h_t = \frac{L\eta \, | t |^2}{(1 - \alpha \, | t |)^2} \leqslant \frac{1}{2},$$

ce qui est équivalent à (25).

En désignant par $z_t$ la solution de l'équation (29), on obtient

$$\| z_t \| = \| z_t - z_0 \| \leqslant \frac{1 - \sqrt{1 - 2h_t}}{h_t} \, \eta \, | t | \leqslant 2\eta \, | t | = B \, | t |,$$

et de (27) il vient

$$| \lambda_t - \lambda_0 | = | \delta | \leqslant | t | \, \| f_0 \| \, \| V \| \, \| x_0 + z_t \| \leqslant A \, | t |.$$

Ceci achève la démonstration du théorème.

Remarque. Il existe un cas particulier important où l'on peut obtenir une expression plus simple pour $\eta$ et $\alpha$. Soient X un espace hilbertien et $U$ un opérateur auto-adjoint. Il est immédiat de vérifier que dans ce cas on peut prendre pour $f_0$ la fonctionnelle définie par l'élément $x_0$, c'est-à-dire

$$f_0(y) = (y, x_0) \quad (y \in \mathrm{X}).$$

Comme

$$\| x_0 \|^2 = f_0(x_0) = 1,$$

il vient

$$\| R(z_0) \| \leqslant \| T \| \, \| Vx_0 - (Vx_0, x_0)\, x_0 \| \leqslant \| T \| \, \| V(x_0) \| \leqslant$$
$$\leqslant \| T \| \, \| V \|,$$
$$\| R'(z_0)(x) \| \leqslant \| T \| \, \| Vx - (Vx, x_0)\, x_0 - (Vx_0, x_0)\, x \| \leqslant$$
$$\leqslant \| T \| \, [\| Vx - (Vx, x_0)\, x_0 \| + \| (Vx_0, x_0)\, x \|] \leqslant$$
$$\leqslant \| T \| \, [\| Vx \| + | (Vx_0, x_0) | \, \| x \|] \leqslant 2 \, \| T \| \, \| V \| \, \| x \|,$$

donc

$$\eta = \| T \| \| V \|, \quad \alpha = 2 \| T \| \| V \|,$$

et la condition de solubilité de l'équation (29) est

$$\frac{2 (\| T \| \| V \| \, | t |)^2}{(1 - 2 \| T \| \| V \| \, | t |)^2} \leqslant \frac{1}{2},$$

ce qui permet de remplacer (25) par

$$\| T \| \| V \| \, | t | \leqslant \frac{1}{4}.$$

L'étude directe de ce cas a conduit M. Gavourine [3] à la condition

$$\| T \| \| V \| \, | t | \leqslant \frac{1}{2}$$

qui est inaméliorable.

3.7. Considérons l'équation différentielle quasi linéaire du deuxième ordre à deux variables indépendantes:

$$A (s, t, u, p, q) \frac{\partial^2 u}{\partial s^2} + B (s, t, u, p, q) \frac{\partial^2 u}{\partial s \, \partial t} + \\ + C (s, t, u, p, q) \frac{\partial^2 u}{\partial t^2} + D (s, t, u, p, q) = 0, \quad (30)$$

où $p = \frac{\partial u}{\partial s}$ et $q = \frac{\partial u}{\partial t}$.

On étudiera l'équation (30) dans un domaine $Q$ et on cherchera la solution qui vérifie une condition aux limites sur le bord $\gamma$ de $Q$.

L'application de la méthode de Newton à ce problème permet d'indiquer la condition d'existence et le domaine de disposition de la solution si l'on connaît une solution approchée $u_0 = u_0 (s, t)$ (cf. Kochelev [1], [2]).

On supposera que l'équation (30) est de type elliptique, plus exactement on admettra que

$$B^2 - 4AC < 0 \quad (31)$$

pour toutes les valeurs des variables ou tout au moins pour celles auxquelles nous aurons affaire.

Pour appliquer la méthode de Newton à l'équation (30) mettons cette dernière sous la forme

$$P (u) = 0, \quad (32)$$

où $P$ est un opérateur d'un espace fonctionnel **U** dans un espace fonctionnel **V**. Ces espaces seront concrétisés plus bas.

Soit $F = F (s, t, u, p, q)$ une fonction quelconque. Soit $F_0$ la fonction obtenue en remplaçant dans $F$ la fonction $u$ par $u_0$, $p$ et $q$ par $p_0 = \frac{\partial u_0}{\partial s}$ et $q_0 = \frac{\partial u_0}{\partial t}$.

Il est immédiat que

$$P'(u_0)(u) = A_0 \frac{\partial^2 u}{\partial s^2} + \left[ \left( \frac{\partial A}{\partial u} \right)_0 u + \left( \frac{\partial A}{\partial p} \right)_0 p + \left( \frac{\partial A}{\partial q} \right)_0 q \right] \frac{\partial^2 u_0}{\partial s^2} +$$
$$+ B_0 \frac{\partial^2 u}{\partial s\, \partial t} + \left[ \left( \frac{\partial B}{\partial u} \right)_0 u + \left( \frac{\partial B}{\partial p} \right)_0 p + \left( \frac{\partial B}{\partial q} \right)_0 q \right] \frac{\partial^2 u_0}{\partial s\, \partial t} +$$
$$+ C_0 \frac{\partial^2 u}{\partial t^2} + \left[ \left( \frac{\partial C}{\partial u} \right)_0 u + \left( \frac{\partial C}{\partial q} \right)_0 p + \left( \frac{\partial C}{\partial q} \right)_0 q \right] \frac{\partial^2 u_0}{\partial t^2} =$$
$$= \left[ A_0 \frac{\partial^2}{\partial s^2} + B_0 \frac{\partial^2}{\partial s\, \partial t} + C_0 \frac{\partial^2}{\partial t^2} \right] u + \ldots \qquad (33)$$

Les termes omis ne contiennent pas de dérivées secondes de $u$. Donc, $P'(u_0)$ est un opérateur différentiel elliptique linéaire et, par suite, chaque pas de la méthode de Newton pour l'équation (32) se ramène à la résolution de l'équation différentielle elliptique linéaire

$$P'(u_n)(u_{n+1}) = P'(u_n)(u_n) - P(u_n) \quad (n = 0, 1, \ldots). \qquad (34)$$

Si l'on se sert du processus modifié, à chaque pas on obtient l'équation elliptique

$$P'(u_0)(u_{n+1}) = P'(u_0)(u_n) - P(u_n) \quad (n = 0, 1, \ldots) \qquad (35)$$

mais avec un premier membre invariant.

La dérivée seconde de l'opérateur $P$, qui nous servira à étudier la convergence de la méthode de Newton, se calcule aussi sans peine. Plus exactement, il est clair que $P''(u)(x, y)$ est une forme bilinéaire de $x$ et $y$ et de leurs dérivées premières et secondes avec des coefficients dépendant de $u$. Il importe de remarquer que cette forme ne contient pas de produit des dérivées secondes de $x$ et $y$.

Penchons-nous maintenant sur les espaces $\mathbf{U}$ et $\mathbf{V}$. Imposons-leur les conditions suivantes:

1) L'opérateur $P$ envoie l'espace $\mathbf{U}$ (ou sa partie $\Omega$) dans l'espace $\mathbf{V}$.

2) L'opérateur $P$ doit être deux fois dérivable et les expressions de ses dérivées sont celles indiquées plus haut.

3) Quel que soit $u \in \Omega$, les coefficients $F_k$ de la forme bilinéaire $P''(u)(x, y)$, traités comme des opérateurs de multiplication, sont des opérateurs linéaires continus de $\mathbf{V}$ dans $\mathbf{V}$ et

$$\| F_k \|_{\mathbf{V}}^{\mathbf{V}} \leqslant M \quad (u \in \Omega). \qquad (36)$$

De façon analogue, si $x \in \mathbf{U}$, la fonction $\frac{\partial^{\alpha+\beta} x}{\partial s^\alpha\, \partial t^\beta}$ $(\alpha, \beta = 0, 1;$ $\alpha + \beta \leqslant 1)$, traitée également comme un opérateur de multiplication, est un opérateur linéaire continu de $\mathbf{V}$ dans $\mathbf{V}$ et

$$\left\| \frac{\partial^{\alpha+\beta} x}{\partial s^\alpha\, \partial t^\beta} \right\|_{\mathbf{V}}^{\mathbf{V}} \leqslant M_1 \| x \|_{\mathbf{U}}. \qquad (37)$$

4) L'opérateur qui associe à l'élément $x \in \mathbf{U}$ sa dérivée $\frac{\partial^{\alpha+\beta} x}{\partial s^\alpha \partial t^\beta}$ ($\alpha, \beta = 0, 1, 2;\ \alpha + \beta \leqslant 2$) est un opérateur linéaire continu de $\mathbf{U}$ dans $\mathbf{V}$.

5) L'opérateur $P'(u_0)$ admet un inverse continu $\Gamma_0 = [P'(u_0)]^{-1}$ de $\mathbf{V}$ dans $\mathbf{U}$.

Les espaces $\mathbf{U}$ et $\mathbf{V}$ remplissant ces conditions seront dits *associés*. Le théorème 1.6 s'applique à un tel couple d'espaces, pourvu que l'approximation initiale $u_0$ soit correctement choisie. En effet, les conditions 3) et 4) nous disent que $\| P''(u) \|$ est uniformément bornée, puisque, par exemple, en vertu de (36) et (37), on a pour le terme de la forme $F \frac{\partial x}{\partial s} \frac{\partial^2 y}{\partial t^2}$

$$\left\| F \frac{\partial x}{\partial s} \frac{\partial^2 y}{\partial t^2} \right\|_{\mathbf{V}} \leqslant \| F \|_{\mathbf{V}}^{\mathbf{V}} \left\| \frac{\partial x}{\partial s} \right\|_{\mathbf{V}}^{\mathbf{V}} \left\| \frac{\partial^2}{\partial t^2} \right\| \| y \|_{\mathbf{U}} \leqslant M \cdot M_1 \left\| \frac{\partial^2}{\partial t^2} \right\| \| x \| \| y \|.$$

On indiquera plus bas quelques couples d'espaces associés. Les conditions 1) à 5) ne seront vérifiées que pour le premier cas, qui est le plus élémentaire.

Considérons l'équation *)

$$P(u) \equiv \pi(u) + R(u) \equiv \Delta u + \frac{\partial}{\partial s}(pH) + \frac{\partial}{\partial t}(qH) = \varphi$$
$$\left(\pi(u) = \Delta u,\ R(u) = \frac{\partial}{\partial s}(pH) + \frac{\partial}{\partial t}(qH)\right) \tag{38}$$

et le problème de Neumann

$$\frac{\partial u}{\partial n} = 0 \tag{39}$$

sur la frontière du carré $Q = [0, \pi;\ 0, \pi]$.

On supposera que la fonction $H$ est bornée et est continue avec toutes les dérivées nécessaires.

La formule de Green nous dit immédiatement que si $u$ vérifie la condition aux limites, alors

$$\iint_Q [P(u)]\, ds\, dt = 0; \tag{40}$$

donc la fonction $\varphi$ doit satisfaire à la condition $\iint_Q \varphi\, ds\, dt = 0$.

---

*) Comme équation de ce type citons l'équation de la surface minimale

$$\frac{\partial}{\partial s} \frac{p}{\sqrt{1+p^2+q^2}} + \frac{\partial}{\partial t} \frac{q}{\sqrt{1+p^2+q^2}} = 0.$$

On rencontre ce type d'équation dans le problème de torsion plastique.

S'agissant de l'équation (38), considérons le couple d'espaces

$$U = \overset{\circ}{W}{}_2^{(3)}, \quad V = \overset{\circ}{W}{}_2^{(1)},$$

où $\overset{\circ}{W}{}_2^{(3)}$ est composé des fonctions $u \in W_2^{(3)}$ vérifiant la condition aux limites (39) et la condition

$$\iint_Q u(s, t)\, ds\, dt = 0. \tag{41}$$

L'espace $\overset{\circ}{W}{}_2^{(1)}$ est formé des fonctions $v \in W_2^{(1)}$ vérifiant la condition

$$\iint_Q v(s, t)\, ds\, dt = 0. \tag{42}$$

Munissons les espaces indiqués des normes respectives

$$\| u \| = \left[ \iint_Q \sum_{\alpha+\beta=3} \left( \frac{\partial^3 u}{\partial s^\alpha \partial t^\beta} \right)^2 ds\, dt \right]^{1/2},$$

$$\| v \| = \left[ \iint_Q \left( \left( \frac{\partial v}{\partial s} \right)^2 + \left( \frac{\partial v}{\partial t} \right)^2 \right) ds\, dt \right]^{1/2} \quad (u \in U,\ v \in V).$$

Vérifions que le couple U, V satisfait bien aux conditions 1) à 5) (pour $\Omega$ on prend une boule de rayon assez grand).

Le théorème d'immersion nous dit que les dérivées secondes de la fonction $u \in U$ sont de puissance $p$-ième sommable quel que soit $p$, donc en particulier

$$\left[ \iint_Q \left( \frac{\partial^2 u}{\partial s^2} \right)^4 ds\, dt \right]^{1/4} \leqslant M_2 \| u \|. \tag{43}$$

Il en est de même pour les autres dérivées secondes. La fonction $u$ et ses dérivées premières $p$ et $q$ sont bornées.

Considérons le terme caractéristique de l'expression de $\| P(u) \|$:

$$\left\{ \iint_Q \left[ \frac{\partial}{\partial s} \left( H \frac{\partial^2 u}{\partial s^2} \right) \right]^2 ds\, dt \right\}^{1/2} \leqslant \left\{ \iint_Q \left( H \frac{\partial^3 u}{\partial s^3} \right)^2 ds\, dt \right\}^{1/2} +$$

$$+ \left\{ \iint_Q \left[ \left( \frac{\partial H}{\partial s} + \frac{\partial H}{\partial u} p + \frac{\partial H}{\partial p} \frac{\partial^2 u}{\partial s^2} + \frac{\partial H}{\partial q} \frac{\partial^2 u}{\partial s\, \partial t} \right) \frac{\partial^2 u}{\partial s^2} \right]^2 ds\, dt \right\}^{1/2} \leqslant$$

$$\leqslant \max | H | \, \| u \| +$$

$$+ \max \left| \frac{\partial H}{\partial q} \right| \left\{ \iint_Q \left( \frac{\partial^2 u}{\partial s\, \partial t} \right)^4 ds\, dt \right\}^{1/4} \left\{ \iint_Q \left( \frac{\partial^2 u}{\partial s^2} \right)^4 ds\, dt \right\}^{1/4} + \ldots$$

Le second terme est fini en vertu de (43); les termes non écrits le sont de toute évidence, donc $P(u) \in V$.

On vérifie par les mêmes raisonnements que $P$ est dérivable et que $P(u_0)$ est de la forme (33). On commence par établir que l'expres-

sion (33) est la dérivée faible, et comme $P'$ est continue, on déduit que $P$ est dérivable. Il en va de même pour la dérivée seconde.

Considérons la fonction $F$ qui est coefficient de la forme bilinéaire $P''(u)(x, y)$. Il est évident que $F$ est la dérivée partielle de la fonction $H$, donc elle est continue et bornée. Pour $v \in \mathbf{V}$ on a

$$\left\{\iint_Q \left[\frac{\partial}{\partial s}(Fv)\right]^2 ds\,dt\right\}^{1/2} =$$

$$=\left\{\iint_Q \left[F\frac{\partial v}{\partial s} + v\left(\frac{\partial F}{\partial s} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial s} + \frac{\partial F}{\partial p}\frac{\partial^2 u}{\partial s^2} + \frac{\partial F}{\partial q}\frac{\partial^2 u}{\partial s\,\partial t}\right)\right]^2 ds\,dt\right\}^{1/2} \leqslant$$

$$\leqslant \max |F|\,\|v\| +$$

$$+\left\{\iint_Q \left[v\left(\frac{\partial F}{\partial s} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial s} + \frac{\partial F}{\partial p}\frac{\partial^2 u}{\partial s^2} + \frac{\partial F}{\partial q}\frac{\partial^2 u}{\partial s\,\partial t}\right)\right]^2 ds\,dt\right\}^{1/2}. \qquad (44)$$

Le théorème d'immersion nous dit que la fonction $v$ est de puissance $p$-ième sommable quel que soit $p$, et, en particulier,

$$\left\{\iint_Q [v(s, t)]^4\,ds\,dt\right\}^{1/4} \leqslant M_3 \|v\|.$$

Utilisons ceci pour majorer l'un des termes de l'expression du milieu de (44). Par exemple

$$\left\{\iint_Q \left[v\frac{\partial F}{\partial p}\frac{\partial^2 u}{\partial s^2}\right]^2 ds\,dt\right\}^{1/2} \leqslant$$

$$\leqslant \max\left|\frac{\partial F}{\partial p}\right| \left\{\iint_Q v^4\,ds\,dt\right\}^{1/4} \left\{\iint_Q \left(\frac{\partial^2 u}{\partial s^2}\right)^4 ds\,dt\right\}^{1/4} \leqslant M_4 \|v\|.$$

De ce qui vient d'être dit, il suit que l'opérateur considéré est un opérateur linéaire continu de $\mathbf{V}$ dans $\mathbf{V}$, et de plus est réalisée (36). La deuxième partie de la condition 3) résulte de ce que la fonction $u \in \mathbf{U}$ est bornée avec ses dérivées $p$ et $q$.

La condition 4) est évidente.

Voyons la condition 5). Pour utiliser les raisonnements de 2.6, il suffit de prouver que l'opérateur de Laplace, traité comme un opérateur de $\mathbf{U}$ dans $\mathbf{V}$, possède un inverse continu. A cet effet, étudions l'équation

$$\Delta u = v \quad (v \in \mathbf{V}). \qquad (45)$$

Développons le second membre en série de cosinus

$$v(s, t) = \sum_{k,\,m=0}^{\infty} a_{km} \cos ks \cos mt.$$

$a_{00} = 0$ en vertu de (42). On voit donc sans peine que l'unique solution de l'équation (45) vérifiant la condition aux limites (39)

est donnée par la série

$$u(s, t) = \sum_{k, m=0}^{\infty} \frac{a_{km}}{k^2+m^2} \cos ks \cos mt.$$

Vérifions que $u \in$ U. En effet, les dérivées de la fonction $v$ étant de carré sommable, on a

$$\sum_{k, m=0}^{\infty} (k^2+m^2)\, a_{km}^2 = M_4 \iint_Q \left[\left(\frac{\partial v}{\partial s}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2\right] ds\, dt = M_4 \| v \|^2.$$

La convergence en moyenne des séries définissant les dérivées troisièmes de la fonction $u$ est alors évidente, par exemple

$$\frac{\partial^3 u(s, t)}{\partial s^2 \partial t} = \sum_{k, m=0}^{\infty} \frac{k^2 m}{k^2+m^2} a_{km} \cos ks \cos mt.$$

Ceci étant

$$\iint_Q \left(\frac{\partial^3 u}{\partial s^2 \partial t}\right)^2 ds\, dt = M_5 \sum_{k, m=0}^{\infty} \frac{k^4 m^2}{(k^2+m^2)^2} a_{km}^2 \leqslant$$

$$\leqslant M_6 \sum_{k, m=0}^{\infty} k^2 a_{km}^2 \leqslant M_7 \| v \|^2.$$

Donc $u \in$ U et $\| u \| \leqslant M_8 \| v \|$; et l'existence de l'opérateur $\Delta^{-1}$ est prouvée.

On peut donc appliquer les théorèmes de convergence de la méthode de Newton. A noter en particulier que si, au lieu de l'équation (38), on considère l'équation

$$\Delta u + \mu \left[\frac{\partial}{\partial s}(pH) + \frac{\partial}{\partial t}(qH)\right] = \varphi,$$

où $\mu$ est un paramètre, alors, la solution existant pour $\mu = 0$, elle existera, d'après le théorème 2.6, pour tous les $\mu$ assez petits.

On remarquera qu'en appliquant le schéma du théorème 2.6, on obtient la solution de l'équation de Poisson à chaque pas.

En conclusion indiquons sans entrer dans les détails deux couples d'espaces associés. Dans le problème de Dirichlet on peut prendre, par exemple, $\mathrm{U} = \dot{\mathrm{W}}_p^{(2)}$, $\mathrm{V} = \mathrm{L}^p$. La vérification des conditions 1) à 4) ne soulève aucune difficulté. Ce qui n'est pas le cas de la condition 5) qui est difficile à établir.

Enfin, pour U on peut prendre l'espace $\mathrm{Lip}^{(2)}\,\alpha$ de toutes les fonctions continûment dérivables, dont les dérivées secondes vérifient la condition de Lipschitz dans le rapport $\alpha$. V sera alors l'espace

Lip $\alpha$. Comme dans le cas précédent, la principale difficulté est la vérification de la condition 5) qui implique l'utilisation de théorèmes fins de la théorie des équations aux dérivées partielles.

D'autres applications de la méthode de Newton sont accessibles dans les travaux de Myssovskikh [2], de Nikolaéva [1].

## § 4. La méthode de Newton dans les espaces réticulés normés

La méthode de Newton, développée au § 1 pour les espaces normés, passe pour une plus vaste classe d'espaces étroitement liés aux lattis vectoriels: la classe des espaces réticulés normés. Les résultats des numéros 4.1 et 4.2 sont dus à L. Kantorovitch [4], [10], [12], de 4.3 à A. Balouev [1]. Notre exposé suit Voulikh [I] dans lequel sont accessibles les démonstrations détaillées des théorèmes énoncés et les références bibliographiques.

4.1. On dit qu'un espace vectoriel réel X est un espace *réticulé normé* par un lattis vectoriel Z si à tout élément $x \in X$ est associé un élément positif $|x| \in Z$, appelé sa *norme abstraite*, vérifiant les axiomes ordinaires:

1) $|x| > 0$ si $x \neq 0$;

2) $|\lambda x| = |\lambda| |x|$;

3) $|x_1 + x_2| \leqslant |x_1| + |x_2|$.

On dira que Z est un lattis vectoriel normant. Un espace normé est un cas particulier d'un espace réticulé normé: dans ce cas le l.v. normant est le $K$-espace des nombres réels. De façon plus générale, tout e.l.c. peut être représenté comme un espace réticulé normé par un $K$-espace $s(T)$, où la puissance de $T$ est égale à celle du système générateur de semi-normes. Enfin, si dans un l.v. X quelconque, pour norme abstraite d'un élément on prend le module de cet élément, alors on fait de X un espace réticulé normé dont le l.v. normant est X lui-même.

Soit X un espace réticulé normé par un l.v. Z. On dira qu'une suite $\{x_n\}$ est $(o\mathbf{Z})$-convergente vers un élément $x \in X$ $(x_n, x \in X)$ et l'on notera $x_n \xrightarrow{(oz)} x$ si $|x_n - x| \xrightarrow{(o\sigma)} 0$ dans X. Une suite $\{x_n\} \subset \subset X$ est dite $(o\mathbf{Z})$-*fondamentale* si dans Z existe une suite $z_m \downarrow 0$ telle que

$$|x_n - x_m| \leqslant z_m \quad \text{pour tous les} \quad n > m.$$

Si toute suite $(o\mathbf{Z})$-fondamentale est $(o\mathbf{Z})$-convergente, on dit que l'espace X est $(o\mathbf{Z})$-*complet*. Un $B$-espace est $(o\mathbf{R}^1)$-complet; un $K$-espace X est $(o\mathbf{X})$-complet.

Les espaces X [Y] à norme mixte introduits dans XI.1.3, tome 1, constituent un important exemple d'espaces réticulés normés.

L'espace X [Y] peut être considéré comme étant normé par le $K$-espace X si la norme abstraite est définie par la formule

$$| K | (t) = \| K (\cdot, t) \|_Y.$$

On laisse au lecteur le soin de vérifier que l'espace X [Y] est ($o$X)-complet.

Introduisons quelques notions indispensables à l'étude de la méthode de Newton.

On supposera dans 4.1 et 4.2 que **X** et **Y** sont des espaces réticulés normés par des $K$-espaces **Z** et **W** respectivement.

Soit $U: X \to Y$ un opérateur linéaire. Un opérateur linéaire positif $U_0: Z \to W$ est *majorant modulaire* d'un opérateur $U$ si

$$| U (x) | \leqslant U_0 (| x |)$$

pour tous les $x \in X$.

Introduisons la notion de dérivée d'un opérateur qui nous servira dans l'étude des équations fonctionnelles.

Soit $P: X \to Y$ un opérateur quelconque. Prenons un élément fixe $x_0 \in X$ et supposons qu'il existe un opérateur linéaire $U: X \to Y$ tel que pour tout $x \in X$ et toute suite de nombres $t_k \to 0$ ($t_k \neq 0$) on ait

$$(o\mathrm{W})\text{-}\lim \frac{P(x_0 + t_k x) - P(x_0)}{t_k} = U(x).$$

On dit alors que l'opérateur linéaire $U$ est la *dérivée* de l'opérateur $P$ au point $x_0$ et l'on écrit

$$U = P'(x_0).$$

Voyons maintenant la notion d'intégrale. Soient $x_0$, $x \in X$. Désignons par $[[x_0, x]]$ l'intervalle reliant les points $x$ et $y$ (cf. II.3.1, tome 1 ; à ne pas confondre avec un intervalle au sens de la structure d'ordre!).

Soit donnée sur l'intervalle $[[x_0, x]]$ une fonction $F(x)$ dont les valeurs sont des opérateurs linéaires de X dans Y.

Soit donnée une suite $\{\sigma_n\}$ de partitions de l'intervalle $[0, 1]$ : $0 = t_0 < t_1 < \ldots < t_n = 1$ telle que $\lambda = \max (t_{k+1} - t_k) \to 0$ et soient fixés $\tau_k \in [t_k, t_{k+1}]$ ($0 \leqslant k \leqslant n - 1$). Posons $\xi_k = (1 - \tau_k) \times x_0 + \tau_k x$. Si dans Y existe la ($o$W)-limite de la suite des sommes intégrales

$$\sum_{k=0}^{n-1} F(\xi_k)(t_{k+1} - t_k),$$

alors cette limite qui ne dépend pas du choix de $\{\sigma_n\}$ et $\{\tau_k\}$ est appelée intégrale de la fonction $F$ et notée $\int_{x_0}^{x} F(x)\, dx$.

Cette définition vaut en particulier pour la dérivée $P'(x)$ de l'opérateur $P: X \to Y$. On dira qu'un opérateur $P$ est justiciable

de la formule de Newton-Leibniz sur un ensemble $\Omega \subset \mathbf{X}$ si $P'(x)$ existe pour tous les $x \in \Omega$ et

$$P(x_0+\Delta x)-P(x_0)=\int_{x_0}^{x_0+\Delta x} P'(x)\,dx$$

pour tous les $\Delta x$ tels que $x_0 + \Delta x \in \Omega$.

Nous n'indiquerons pas les conditions de validité de la formule de Newton-Leibniz, en laissant au lecteur le soin de formuler une condition suffisante à partir de la démonstration de XVII.1.7 (cf. Voulikh [I], page 372).

4.2. Supposons maintenant que $\mathbf{X}$ est $(o\mathbf{Z})$-complet. Soit l'équation

$$P(x) = 0, \tag{1}$$

où $P$ est un opérateur de $\mathbf{X}$ dans $\mathbf{Y}$. Fixons $x_0 \in \mathbf{X}$. Comme dans 1.1, pour résoudre l'équation (1) nous pouvons utiliser la *méthode de Newton*

$$x'_{n+1} = x'_n - [P'(x_n)]^{-1}(P(x'_n)) \quad (n = 0, 1, \ldots; \quad x'_0 = x_0)$$

et la *méthode de Newton modifiée*

$$x_{n+1} = x_n - [P'(x_0)]^{-1}(P(x_n)) \quad (n = 0, 1, \ldots). \tag{2}$$

Nous nous limitons à une étude brève de la méthode modifiée de Newton. Posons $\Gamma_0 = [P'(x_0)]^{-1}$ (la seule condition imposée à $\Gamma_0$ est d'être un opérateur linéaire défini sur $\mathbf{Y}$ tout entier). Outre l'équation (1) considérons l'équation

$$Q(z) = 0, \tag{3}$$

où $Q$ est un opérateur de $\mathbf{Z}$ dans $\mathbf{W}$, qui sera équation *majorante*. Soit donné dans $\mathbf{Z}$ l'$(o)$-intervalle $[z_0, z']$;

$$D = \{x \in \mathbf{X} : |x - x_0| \leqslant z' - z_0\}.$$

THÉORÈME 1. *Supposons que les opérateurs $P$ et $Q$ vérifient les conditions suivantes:*

1) *l'opérateur $P$ est justiciable de la formule de Newton-Leibniz sur l'ensemble $D$;*

2) *l'opérateur $Q$ est justiciable de la formule de Newton-Leibniz sur l'intervalle $[z_0, z']$;*

3) *si $\{z_n\}_{n=1}^{\infty} \subset [z_0, z']$ et $z_n \xrightarrow{(o)} z$, alors $Q(z_n) \to Q(z)$;*

4) *existent les opérateurs inverses: $\Gamma_0 = [P'(x_0)]^{-1}$ défini sur $\mathbf{Y}$ tout entier, et $\Delta_0 = [Q'(z_0)]^{-1}$ défini sur $\mathbf{W}$ tout entier, et de plus $\Delta_0 \in \widetilde{\mathbf{L}}_{nn}(\mathbf{W}, \mathbf{Z})$ et $\Delta_0$ est majorant modulaire pour $\Gamma_0$;*

5) $|\Gamma_0 P(x_0)| \leqslant -\Delta_0 Q(z_0)$;

6) *si $x \in D$ et $z \in [z_0, z']$ sont tels que $|x - x_0| \leqslant z - z_0$, alors l'opérateur $I - \Delta_0 Q'(z)$ est majorant modulaire pour $I - \Gamma_0 P'(x)$;*

7) *l'équation* (3) *admet une solution* $\tilde{z}$ *sur l'intervalle* $[z_0, z']$.

*L'équation* (1) *admet alors une solution* $\tilde{x}$ *vérifiant la condition* $|\tilde{x} - x_0| \leqslant \tilde{z} - z_0$, *vers laquelle* (*oZ*)-*converge le processus modifié de Newton d'origine* $x_0$.

Si dans les conditions du théorème 1, l'équation (3) admet une solution unique sur l'intervalle $[z_0, z']$, alors l'équation (1) admet une solution unique sur l'ensemble $D$.

La démonstration de ces résultats figure dans Voulikh [1] (théorèmes XII.5.1 et XII.5.2).

**4.3.** Si l'équation fonctionnelle est résolue sur un espace ordonné, on conçoit qu'il faille chercher deux suites monotones convergeant vers cette solution de divers côtés. Les résultats exposés peuvent être interprétés comme une généralisation abstraite de la *méthode de S. Tchapliguine* de la théorie des équations différentielles ordinaires.

Soit l'équation

$$U(x) = 0, \tag{4}$$

où $U$ est un opérateur d'un $K$-espace X dans un l.v. Y.

LEMME 1. *Soient réalisées les conditions suivantes:*

1) *il existe* $x_0, x' \in$ X *tels que*

$$x_0 \leqslant x'_0 \quad \text{et} \quad U(x_0) \leqslant 0 \leqslant U(x'_0);$$

2) *il existe un opérateur linéaire* $T$ *de* X *sur* Y *tout entier tel que si* $x_0 \leqslant x \leqslant x'_0$, *alors*

$$U(x'_0) + T(x - x'_0) \leqslant U(x) \leqslant U(x_0) + T(x - x_0); \tag{5}$$

3) *l'opérateur* $T$ *admet un inverse positif* $T^{-1}$.

*Alors*

a) *les éléments* $x_1$ *et* $x'_1$ *définis par*

$$x_1 = x_0 - T^{-1}U(x_0), \quad x'_1 = x'_0 - T^{-1}U(x'_0)$$

*vérifient les inégalités*

$$x_0 \leqslant x_1 \leqslant x'_1 \leqslant x'_0, \quad U(x_1) \leqslant 0 \leqslant U(x'_1); \tag{6}$$

b) *si l'équation* (4) *admet une solution* $\tilde{x}$ *sur l'*(*o*)-*intervalle* $[x_0, x'_0]$, *alors* $x_1 \leqslant \tilde{x} \leqslant x'_1$.

DÉMONSTRATION. Comme $U(x_0) \leqslant 0$ et l'opérateur $T^{-1} \geqslant 0$, il vient $T^{-1}U(x_0) \leqslant 0$, d'où $x_1 \geqslant x_0$. De façon analogue, $x'_1 \leqslant x'_0$.

Par ailleurs

$$x'_0 - x_1 = x'_0 - x_0 + T^{-1}U(x_0) = T^{-1}T(x'_0 - x) + T^{-1}U(x_0) = \\ = T^{-1}[T(x'_0 - x) + U(x_0)].$$

Le second membre de (5) et la condition 1) entraînent

$$T(x'_0 - x_0) + U(x_0) \geqslant U(x'_0) \geqslant 0,$$

d'où $x'_0 - x_1 \geqslant 0$. Donc $x_0 \leqslant x_1 \leqslant x'_0$ et l'inégalité (5) est réalisée pour $x = x_1$. D'où

$$U(x_1) \leqslant U(x_0) + T(x_1 - x_0) = U(x_0) - TT^{-1}U(x_0) = 0$$

et

$$U(x'_0) + T(x_1 - x'_0) \leqslant U(x_1) \leqslant 0. \tag{7}$$

D'autre part

$$x'_1 - x_1 = x'_0 - T^{-1}U(x'_0) - x_1 = T^{-1}T(x'_0 - x_1) - T^{-1}U(x'_0) =$$
$$= -T^{-1}[U(x'_0) + T(x_1 - x'_0)],$$

d'où, en vertu de (7), $x'_1 - x_1 \geqslant 0$, c'est-à-dire $x_1 \leqslant x'_1$.

En portant $x = x'_1$ dans le premier membre de (5), on obtient

$$U(x'_1) \geqslant U(x'_0) + T(x'_1 - x'_0) = U(x'_0) - TT^{-1}U(x'_0) = 0.$$

Ce qui prouve les inégalités (6). Reste à vérifier la proposition b).

Soient $x_0 \leqslant \tilde{x} \leqslant x'_0$ et $U(\tilde{x}) = 0$. En vertu de (5)

$$\tilde{x} - x_1 = \tilde{x} - x_0 + T^{-1}U(x_0) = T^{-1}[U(x_0) + T(\tilde{x} - x_0)] \geqslant T^{-1}U(\tilde{x}) = 0,$$

d'où $\tilde{x} \geqslant x_1$. On établit de façon analogue que $\tilde{x} \leqslant x'_1$. Et le lemme est prouvé.

THÉORÈME 2. *Soient réalisées les conditions suivantes:*

1) $x_0 \leqslant x'_0$ *et* $Ux_0 \leqslant 0 \leqslant Ux'_0$;

2 *si* $\{\tilde{x}_n\}_{n=1}^{\infty} \subset [x_0, x'_0]$ *et* $\tilde{x}_n \xrightarrow{(o)} x$, *alors* $U(\tilde{x}_n) \xrightarrow{(o\sigma)} U(x)$;

3) *l'opérateur* $U$ *est justiciable de la formule de Newton-Leibniz sur l'intervalle* $[x_0, x'_0]$;

4) *il existe un opérateur linéaire* $T$ *de* $\mathbf{X}$ *sur* $\mathbf{Y}$ *tout entier tel que pour tous les* $x \in [x_0, x'_0]$ *on ait* $[U'(x)](h) \leqslant T(h)$ *pour tout* $h \in \mathbf{X}_+$;

5) *l'opérateur* $T$ *admet un opérateur inverse positif* ($o$)-*continu* $T^{-1}$.

*Alors l'ensemble des solutions de l'équation* (4) *appartenant à l'intervalle* $[x_0, x'_0]$ *n'est pas vide et admet un minimum* $x^*$ *et un maximum* $x^{**}$. *Si, de plus, les éléments* $x_n$ *et* $x'_n$ *sont définis par les formules récurrentielles*

$$x_n = x_{n-1} - T^{-1}U(x_{n-1}), \quad x'_n = x'_{n-1} - T^{-1}U(x'_{n-1}) \quad (n \in \mathbb{N}), \tag{8}$$

*alors* $x_n \uparrow x^*$ *et* $x'_n \downarrow x^{**}$.

DÉMONSTRATION. Montrons que les opérateurs $U$ et $T$ vérifient les conditions du lemme 1 sur tout intervalle $[x_n, x'_n]$.

Prouvons que pour tous $\bar{x}$, $x$ et $\bar{x}'$ tels que

$$x_0 \leqslant \bar{x} \leqslant x \leqslant \bar{x}' \leqslant x'_0,$$

on a l'inégalité analogue à (5):

$$U(\bar{x}') + T(x - \bar{x}') \leqslant U(x) \leqslant U(\bar{x}) + T(x - \bar{x}), \tag{9}$$

En effet, la condition 3) et la définition de l'intégrale nous donnent

$$U(x)-U(\bar{x})=\int_{\bar{x}}^{x} U'(x)\,dx \leqslant \int_{\bar{x}}^{x} T\,dx = T(x-\bar{x}),$$

$$U(\bar{x}')-U(x)=\int_{x}^{\bar{x}'} U'(x)\,dx \leqslant \int_{x}^{\bar{x}'} T\,dx = T(\bar{x}'-x),$$

d'où l'inégalité (9).

En se servant de (9) et de la condition 1) on s'assure par récurrence que les conditions du lemme 1 sont remplies pour tous $x_n$ et $x'_n$, donc

$$x_0 \leqslant x_1 \leqslant \ldots \leqslant x_n \leqslant \ldots \leqslant x'_n \leqslant \ldots \leqslant x'_1 \leqslant x'_0.$$

L'espace X étant un $K$-espace, il existe $x^* = \sup x_n$, $x^{**} = \inf x'_n$ et de plus $x_n \uparrow x^*$, $x'_n \downarrow x^{**}$. En passant à la $(o)$-limite dans les formules (8) et en tenant compte de 2) et de 5), on obtient

$$T^{-1}U(x^*) = T^{-1}U(x^{**}) = 0,$$

d'où

$$U(x^*) = U(x^{**}) = 0,$$

c'est-à-dire $x^*$ et $x^{**}$ sont solutions de l'équation (4). De plus, le lemme 1 nous dit que toute solution $\tilde{x}$ de l'équation (4), contenue dans l'intervalle $[x_0, x'_0]$, vérifie $x_n \leqslant \tilde{x} \leqslant x'_n$ pour tous les $n$, donc $x^* \leqslant \tilde{x} \leqslant x^{**}$. Ce qui achève la démonstration du théorème.

# ANNEXE

## Dualité entre espaces vectoriels

Nous étudions en annexe la dualité entre espaces vectoriels.

Si dans l'espace vectoriel L (X) de toutes les fonctionnelles linéaires définies sur un espace vectoriel donné X, on considère une partie assez riche, il existe alors dans X une multitude de topologies localement convexes pour chacune desquelles ces fonctionnelles sont continues. Si, de plus, l'on exige que seules les fonctionnelles soient continues pour les topologies envisagées, alors la description complète de ces topologies constitue le théorème de Mackey-Arens dont la démonstration est le but de l'annexe.

Seul le paragraphe 2 est consacré à la dualité proprement dite. Dans le paragraphe 1 on expose des faits supplémentaires relatifs à la théorie des correspondances et à la théorie des fermetures sur des ensembles ordonnés. En effet, quand on étudie des espaces vectoriels topologiques on est conduit à analyser les liens existant entre la structure topologique et la structure vectorielle qui est de nature foncièrement algébrique. Il se trouve en outre que la fermeture topologique des ensembles qui est entièrement caractérisée par les axiomes de Kuratowski, et les diverses enveloppes des ensembles (telles les enveloppes absolument convexes, affines, etc.) dont il est fait un large usage, présentent de nombreuses affinités. Beaucoup de faits les concernant, eux et leurs liaisons, peuvent être obtenus selon un schéma unique comme cas particuliers de la théorie générale des fermetures sur des ensembles ordonnés et notamment de la théorie des fermetures introduite par E. Moore. Les ensembles fermés (pour une topologie), les ensembles absolument convexes, les ensembles linéaires, etc., sont traités comme des « points » fixes des opérateurs de fermeture correspondants. La continuité des opérations algébriques peut aussi être exprimée en termes d'opérateurs de fermeture.

Il nous a paru utile d'inclure dans cette annexe certains faits généraux sur les liaisons de Galois engendrées par des correspondances entre deux ensembles. Ces liaisons, appelées polaires et polaires inverses, servent à étudier la dualité entre espaces vectoriels.

Le texte de l'annexe est dû dans son ensemble à G. Akilov. Les auteurs tiennent à exprimer leur reconnaissance à G. Syrkine qui a travaillé à cette annexe en améliorant le texte du § 1 (en particulier des nos 1.1, 1.2) et en rédigeant entièrement les nos 1.3, 1.4 (à l'exception du théorème 11) et 1.7.

## § 1. Correspondances. Fermetures sur des ensembles ordonnés

1.1. Pour exposer en détail les problèmes mentionnés plus haut nous aurons besoin de la notion de correspondance ou application multivoque.

Nous appellerons *relation binaire* tout ensemble dont les éléments sont des couples ordonnés. *)

Si $F$ est une relation binaire, on dit qu'un couple ordonné $(x, y)$ vérifie $F$ et on note $xFy$ (ou encore $F(x, y)$) si et seulement si $(x, y) \in F$. Il est clair que toute relation binaire $F$ est confondue avec l'ensemble de tous les couples ordonnés $(x, y)$ vérifiant cette relation: $F = \{(x, y): xFy\}$. Signalons que nous appelons relation binaire ce que Bourbaki appelle *graphe* dans la « Théorie des ensembles ».

Soit $F$ une relation binaire. L'ensemble des premières composantes des couples de $F$ s'appelle *domaine de définition* ou encore *première projection* de $F$ et se note $\mathscr{D}(F)$ ou $\mathrm{Pr}_1(F)$. L'ensemble des deuxièmes composantes des couples de $F$ s'appelle *domaine des valeurs* ou encore *deuxième projection* et se note $\mathscr{R}(F)$ ou $\mathrm{Pr}_2(F)$. Il est évident que l'inclusion $F \subset A \times B$ est équivalente à la conjonction des inclusions $\mathscr{D}(F) \subset A$ et $\mathscr{R}(F) \subset B$ quels que soient les ensembles $A$ et $B$ et la relation binaire $F$. L'ensemble $\mathscr{D}(F) \cup \mathscr{R}(F)$ s'appelle *corps de la relation binaire* $F$ et se note $\mathscr{F}\ell d(F)$. Il est manifeste que le corps de la relation binaire $F$ est confondu avec le plus petit (au sens de $\subset$) des ensembles $A$ tels que $F \subset A \times A$. Tout triplet **) ordonné $(X, Y, F)$, où $F \subset X \times Y$, s'appelle *correspondance* de domaine de départ $X$, de domaine d'arrivée $Y$ et de graphe $F$.

---

*) On introduit la notion de couple ordonné comme une notion primaire à l'aide de l'axiome:

$$\forall x_1, x_2, y_1, y_2 \, [(x_1, x_2) = (y_1, y_2) \Rightarrow x_1 = y_1 \,\&\, x_2 = y_2]$$

comme dans Bourbaki par exemple. On peut encore adopter la définition « explicite » d'un couple ordonné suivant Kuratowski:

$$(x_1, x_2) = \{\{x_1\}, \{x_1, x_2\}\}$$

qui est généralement utilisée en logique mathématique. Il est immédiat de vérifier que deux couples ordonnés au sens de Kuratowski sont égaux si et seulement si leurs composantes le sont.

**) On peut définir la notion de triplet ordonné en termes de couples ordonnés, par exemple ainsi: $(x, y, z) = ((x, y), z)$.

On appelle *inverse* d'une relation binaire donnée $F$ et on note $F^{-1}$ une relation composée de tous les couples $(x, y)$ tels que $(y, x) \in F$. La correspondance $(Y, X, F^{-1})$ est dite *inverse* ou *réciproque* de $(X, Y, F)$ et notée $(X, Y, F)^{-1}$. Il est immédiat que $(F^{-1})^{-1} = F$ et $\mathscr{R}(F) = \mathscr{D}(F^{-1})$ pour toute relation (ou toute correspondance) $F$.

Soient $\Gamma = (X, Y, F)$ une correspondance, $A$ et $B$ des ensembles arbitraires. Par $(A \dashv \Gamma \vdash B)$ on désigne alors la correspondance $(X, Y, F \cap (A \times B))$ et on l'appelle *trace* (ou *restriction*) de $\Gamma$ sur (resp. à) l'ensemble $A \times B$. En particulier, si $A = X$, la correspondance $(X \dashv \Gamma \vdash B)$ est notée $(\Gamma \vdash B)$; si $B = Y$, la correspondance $(A \dashv \Gamma \vdash Y)$ est notée $(A \dashv \Gamma)$ ou encore $\Gamma \mid A$ et s'appelle *restriction* de $\Gamma$ à l'ensemble $A$. Ces notations et notions étant valables aussi bien pour les correspondances que pour les relations qui en sont les graphes, nous les formulerons généralement soit pour les correspondances soit pour les relations. D'autre part, par abus de langage nous identifierons parfois les correspondances à leurs graphes si l'on sait clairement de quels domaines de départ et d'arrivée il est question.

On appelle *image d'un ensemble M par une correspondance* $\Gamma = (X, Y, F)$ et on note $\Gamma[M]$ ou $F[M]$ l'ensemble de tous les éléments $y$ pour lesquels existent des éléments $x \in M$ tels que $(x, y) \in F$. Il est évident que

$$\Gamma[M] = F[M] = \{y : (\exists x \in M)\,[(x, y) \in F]\} = \mathscr{R}(\Gamma \mid M).$$

L'image d'un ensemble $M$ par la correspondance inverse de $\Gamma = (X, Y, F)$, c'est-à-dire l'ensemble $\Gamma^{-1}[M] = F^{-1}[M]$, s'appelle *image réciproque de M par* $\Gamma$. Il est immédiat que

$$\Gamma^{-1}[M] = \{x : (\exists y \in M)\,[(x, y) \in F]\} = \mathscr{R}(\Gamma^{-1} \mid M) = \mathscr{D}(\Gamma \mid M).$$

L'opération $\Gamma[\cdot]$ de passage à l'image d'une partie d'un ensemble $X$ par $\Gamma = (X, Y, F)$ préserve la réunion des ensembles: pour tout ensemble $\mathfrak{A}$ de parties de $X$

$$\Gamma\Big[\bigcup_{A \in \mathfrak{A}} A\Big] = \bigcup_{A \in \mathfrak{A}} \Gamma[A]. \tag{1}$$

Signalons que cette opération ne préserve pas en général l'intersection des ensembles.

Toute propriété de l'application $\Gamma[\cdot]$ de l'ensemble $\mathfrak{P}(X)$ des parties de $X$ dans l'ensemble $\mathfrak{P}(Y)$ des parties de $Y$, ayant lieu pour toute correspondance $\Gamma$ de $X$ dans $Y$, peut être déduite comme une conséquence purement logique de l'identité (1), puisqu'il est évident que toute application $\varphi$ de $\mathfrak{P}(X)$ dans $\mathfrak{P}(Y)$, vérifiant (1) (c'est-à-dire préservant la réunion des ensembles), se représente de façon unique par une opération de passage à l'image par une cor-

respondance (unique) $\Gamma$ entre $X$ et $Y$, c'est-à-dire $\varphi = \Gamma [\cdot]$. Au lieu de $\Gamma [M]$ on se servira parfois de la notation plus brève $\Gamma M$.

On dit qu'une relation binaire $F$ est *univoque* ou *fonctionnelle* (en le deuxième argument) si :

$$\forall x, y_1, y_2 [(x, y_1) \in F \& (x, y_2) \in F \Rightarrow y_1 = y_2]. \qquad (2)$$

Les correspondances univoques ou fonctions peuvent désormais être définies comme des correspondances dont les graphes sont fonctionnels (en le deuxième argument) ou ce qui revient au même comme des correspondances par lesquelles l'image de chaque ensemble à un élément sera soit un ensemble à un élément, soit un ensemble vide.

Si $\Gamma$ est une correspondance univoque et $\Gamma [\{x\}] = \{y\}$, alors l'élément $y$ s'appelle *valeur* de la correspondance $\Gamma$ en $x$ et se note $\Gamma (x)$ ; de sorte que $\Gamma [\{x\}] = \{\Gamma (x)\}$ pour tous les $x \in \mathscr{D} (\Gamma)$. A noter que pour toute correspondance univoque $\Gamma$ de $X$ vers $Y$ et pour tous ensembles $A$ et $B$ l'image $\Gamma [A]$ est confondue avec l'ensemble des éléments de la forme $\Gamma (x)$ lorsque $x$ parcourt l'ensemble $A \cap \mathscr{D} (\Gamma)$ ; l'image réciproque $\Gamma^{-1} [B]$ est confondue avec l'ensemble $\{x \in \in \mathscr{D} (\Gamma) : \Gamma (x) \in B\}$.

L'opération $\Gamma^{-1} [\cdot]$ de passage aux images réciproques des parties d'un ensemble $Y$ par une correspondance $\Gamma$ de $X$ vers $Y$ préserve les intersections non vides : pour tout ensemble non vide $\mathfrak{A}$ de parties de $Y$ on a

$$\Gamma^{-1} [\bigcap_{A \in \mathfrak{A}} A] = \bigcap_{A \in \mathfrak{A}} \Gamma^{-1} [A]. \qquad (3)$$

Toutes les propriétés de l'application $\Gamma^{-1} [\cdot]$ de l'ensemble des parties $\mathfrak{P} (Y)$ dans l'ensemble $\mathfrak{P} (X)$ des parties de $X$ qui ont lieu pour toutes les correspondances univoques $\Gamma$ de $X$ vers $Y$ sont des conséquences purement logiques des identités (1) et (3), puisque toute application de $\mathfrak{P} (Y)$ dans $\mathfrak{P} (X)$ préservant les réunions et les intersections non vides se représente de façon unique sous forme de l'opération $\Gamma^{-1} [\cdot]$, où $\Gamma$ est une correspondance univoque de $X$ vers $Y$.

La notion d'application que nous avons définie de manière peu rigoureuse comme un procédé associant à chaque élément d'un ensemble $X$ un élément d'un ensemble $Y$ peut maintenant être définie de façon plus formelle comme une correspondance univoque $\Gamma = = (X, Y, F)$ définie sur le domaine de départ tout entier : $\mathscr{D} (\Gamma) = = \mathscr{D} (F) = X$.

Le graphe de toute correspondance univoque $\Gamma$ est manifestement l'ensemble

$$\{(x, y) : y = \Gamma (x)\} = \{(x, \Gamma (x)) : x \in \mathscr{D} (\Gamma)\}.$$

On appelle *composition* de deux relations binaires $F$ et $G$ la relation binaire

$$\{(x, z) : \exists y [(x, y) \in F \& (y, z) \in G]\}$$

que l'on note $G \circ F$ *). On appelle *composée des correspondances* $(X, Y, F)$ et $(Z, V, G)$ la correspondance $(X, V, G \circ F)$ qui est notée $(Z, V, G) \circ (X, Y, F)$. Dans certains cas (par exemple en théorie des catégories) la composée est considérée comme définie seulement si $Y = Z$. Les propriétés suivantes de la composition sont une conséquence immédiate de la définition:

$$(G \circ F)[A] = G[F[A]],$$

$$\mathscr{R}(G \circ F) = G[\mathscr{R}(F)],$$

$$\mathscr{D}(G \circ F) = F^{-1}[\mathscr{D}(G)],$$

$$\mathscr{R}(G \circ F) \subset \mathscr{R}(G), \quad \mathscr{D}(G \circ F) \subset \mathscr{D}(F),$$

$F$ et $G$ étant des correspondances (ou des relations) quelconques, $A$ un ensemble arbitraire.

La composition est associative, c'est-à-dire $H \circ (G \circ F) = (H \circ G) \circ F$, ce qui nous permet d'écrire $H \circ G \circ F$ pour $H \circ (G \circ F)$ et $(H \circ G) \circ F$. L'associativité de la composition découle du fait que chacune des deux conditions

$$(x, u) \in H \circ (G \circ F) \quad \text{et} \quad (x, u) \in (H \circ G) \circ F$$

est équivalente à la condition suivante:

$$\exists y, z\, [(x, y) \in F \,\&\, (y, z) \in G \,\&\, (z, u) \in H].$$

On peut du reste établir l'associativité à partir des propriétés de la composition signalées plus haut de la manière suivante. Commençons par remarquer que d'une façon générale la condition de coïncidence des correspondances $\Gamma_1$ et $\Gamma_2$ de $X$ vers $Y$ équivaut à la coïncidence des opérations de passage à l'image $\Gamma_1[\cdot]$ et $\Gamma_2[\cdot]$, c'est-à-dire à la condition

$$(\forall A \subset X)\, [\Gamma_1[A] = \Gamma_2[A]].$$

Or, cette condition équivaut à son tour à la condition plus simple

$$(\forall x \in X)\, [\Gamma_1\{x\} = \Gamma_2\{x\}].$$

Reste maintenant à remarquer que

$$(H \circ (G \circ F))[A] = H[(G \circ F)[A]] = H[G[F[A]]] =$$
$$= (H \circ G)[F[A]] = ((H \circ G) \circ F)[A], \quad \forall A \subset X.$$

Il est évident que la composée de trois correspondances est:

$$H \circ G \circ F = \bigcup_{(y, z) \in G} (F^{-1}\{y\} \times H\{z\}).$$

Soit $A$ un ensemble arbitraire. La relation

$$I_A = \{(x, y) \in A^2 : x = y\} = \{(x, x) : x \in A\}$$

*) On écrit encore $F \circ G$.

s'appelle *relation d'identité* dans l'ensemble $A$. Il est clair que $\mathscr{D}(I_A) = \mathscr{R}(I_A) = A$, que $I_A$ est univoque et que de plus $I_A(x) = x$ pour tout $x \in A$. Si l'ensemble $A$ est contenu dans un ensemble $B$, la relation $I_A$ traitée comme une correspondance de $A$ vers $B$ (ou ce qui revient au même comme une application de $A$ dans $B$) s'appelle *injection canonique* de $A$ dans $B$. Si $F$ est une correspondance, on vérifie immédiatement que

$$F^{-1} \circ F \supset I_{\mathscr{D}(F)}, \quad F \circ F^{-1} \supset I_{\mathscr{R}(F)}$$

et de plus

$$(A \mid F \mid B) = I_B \circ F \circ I_A, \quad \forall A, B.$$

Si la correspondance $F$ est univoque, la deuxième de ces trois relations se transforme en égalité: pour tout $y \in \mathscr{R}(F)$ l'appartenance $x \in F^{-1}\{y\}$ équivaut à $x \in \mathscr{D}(F)$ et $F(x) = y$; donc, l'image $F[F^{-1}\{y\}]$ composée, en vertu de l'univocité de $F$, de tous les éléments de la forme $F(x)$ $(x \in F^{-1}\{y\})$ est confondue avec l'ensemble $\{y\} = I_{\mathscr{R}(F)}\{y.\}$

Formulons maintenant un critère d'univocité des correspondances.

I. *Dire qu'une correspondance $F$ de $X$ vers $Y$ est univoque revient à dire que $F \circ F^{-1} \subset I_Y$ ou encore qu'il existe une correspondance $\Phi$ de $Y$ vers $X$ telle que $\mathscr{D}(F) \subset \mathscr{R}(\Phi)$ et $F \circ \Phi \subset I_{\mathscr{R}(F)}$. Ces deux inclusions entraînent en outre $\Phi^{-1} \supset F$.*

Ce critère ne sera pas prouvé, car trivial.

En utilisant cette proposition pour le cas où $f = F^{-1}$ est une application, on déduit sans peine le critère suivant d'injectivité.

II. *Une application $f$ d'un ensemble $X$ dans un ensemble $Y$ est injective si et seulement si existe une correspondance $\varphi$ de $Y$ vers $X$ telle que $\varphi \circ f = I_{\mathscr{D}(F)} = I_X$. De ces égalités il résulte en particulier que $\varphi \supset f^{-1}$.*

Toute application $\varphi$ de $Y$ dans $X$ vérifiant la condition $\varphi \circ f = I_{\mathscr{D}(f)}$ est dite *inverse à gauche* de $f$.

On dit qu'une correspondance $\Gamma$ de $X$ vers $X$ est une *involution* dans $X$ si $\mathscr{D}(\Gamma) = \mathscr{R}(\Gamma) = X$ et $\Gamma \circ \Gamma = I_X$. L'univocité de l'involution $\Gamma$ résulte de la proposition I, donc l'injectivité de $\Gamma$ découle de la proposition II. On a $\Gamma^{-1} = \Gamma$ puisque $\Gamma \subset \Gamma^{-1}$ d'après la proposition I, donc $\Gamma^{-1} \subset \Gamma$.

Comme exemple d'involution on a l'application du plan complexe dans lui-même, qui à tout nombre complexe $z$ associe son conjugué complexe $\bar{z}$. On obtient un autre exemple d'involution si à tout élément $x$ d'un espace vectoriel donné on associe l'élément $-x$. Enfin en associant à toute partie $A$ d'un ensemble $T$ le complémentaire $T \setminus A$ (relativement à $T$) on a un troisième exemple d'involution dans l'ensemble $\mathfrak{P}(T)$ de toutes les parties de l'ensemble $T$.

On dit qu'une correspondance $P$ est *idempotente* si $P \circ P = P$. L'idempotence d'une application $P$ est équivalente à la relation $(\forall x \in \mathscr{D}(P))\,[P(P(x)) = P(x)]$ qui est l'expression algébrique d'une condition disant que tout élément de l'ensemble de valeurs de $P$ est un point fixe de cette application: $(\forall y \in \mathscr{R}(P))\,[P(y) = y]$. (Par point fixe d'une correspondance $F$ on comprend tout point $y$ tel que $(y, y) \in F$.) Pour l'application $F$ il est équivalent de dire que $F(y) = y$. L'ensemble de tous les points fixes de $F$ se désigne par Fix $(F)$. Vu que chaque point fixe d'une correspondance appartient toujours à l'ensemble de valeurs de $F$ (ainsi qu'à son ensemble de définition), nous pouvons « renforcer » la remarque faite plus haut sur l'idempotence d'une application. Plus exactement, pour qu'une application $P$ soit idempotente, il est nécessaire et suffisant que l'ensemble de ses valeurs soit confondu avec celui de ses points fixes: $\mathscr{R}(P) = \mathrm{Fix}\,(P)$. Dans le langage géométrique, dire qu'une application $P$ est idempotente revient à dire que cette application « projette » son domaine de définition dans son domaine de valeurs en laissant fixes ses valeurs et elles seules. Pour cette raison, les applications idempotentes seront encore appelées projecteurs.

Il existe un critère simple d'idempotence qui indique comment obtenir en général des applications idempotentes.

*Pour qu'une application $P$ soit idempotente (c'est-à-dire un projecteur) il est nécessaire et suffisant qu'elle se représente par la composée $P = f \circ \varphi$, où $f$ et $\varphi$ sont des applications telles que $\varphi \circ f = I_{\mathscr{D}(f)}$.*

◀ La nécessité découle du fait que si $P$ est idempotente, pour $\varphi$ et $f$ on peut manifestement prendre $P$ et $I_{\mathscr{R}(P)}$.

La suffisance résulte de ce que

$$P \circ P = (f \circ \varphi) \circ (f \circ \varphi) = f \circ (\varphi \circ f) \circ \varphi = f \circ I_{\mathscr{D}(f)} \circ \varphi = f \circ \varphi = P. \ ▶$$

**1.2.** Dans ce numéro nous exposons les principaux faits qui concernent les fermetures sur des ensembles ordonnés, c'est-à-dire des opérations idempotentes, croissantes, dilatantes sur ces ensembles. Nous aborderons les problèmes de caractérisation des ensembles de points fixes des fermetures (dits éléments fermés) et établirons les liens existant entre les bornes pour un système d'éléments fermés et les bornes pour l'ensemble ordonné initial sur lequel est définie la fermeture.

Introduisons quelques notations liées aux ensembles ordonnés. Soit donné un ensemble ordonné $\mathbf{X} = (X, \leqslant)$. Pour deux sous-ensembles quelconques $A$ et $M$ de $X$ (plus exactement pour des sous-ensembles du support $X$ de $\mathbf{X}$), désignons par $\beta^-_M(A)$ l'ensemble des bornes inférieures (ou minorants) de $A$ qui sont éléments de l'ensemble $M$: $\beta^-_M(A) = \{x \in M : (\forall y \in A)\ (x \leqslant y)\}$. Il est évident que $\beta^-_M(A) = \beta^-_X(A) \cap M$. De façon duale, par $\beta^+_M(A)$ on désignera l'ensemble: $\{x \in M : (\forall y \in A)\,(y \leqslant x)\}$. Ici aussi $\beta^+_M(A) = \beta^+_X(A) \cap$

$\bigcap M$. Au lieu de $\beta_X^-$ et $\beta_X^+$ on écrira respectivement $\beta^-$ et $\beta^+$. Si $x \in X$, l'ensemble $\{y \in M : x \leqslant y\}$ sera appelé *coupe inférieure de M suivant x* et noté $M_x$ ou encore $[x, \rightarrow)_M$. De façon duale, l'ensemble $\{y \in M : y \leqslant \leqslant x\}$ sera appelé *coupe supérieure de M suivant x* et noté $M^x$ ou encore $(\leftarrow, x]_M$. Si $M = X$, au lieu de $X_x$ et $[x, \rightarrow)_X$ on écrira $[x, \rightarrow)$ et au lieu de $X^x$ et $(\leftarrow, x]_X$, $(\leftarrow, x]$. On désignera par Min $A$ le plus petit élément ou élément minimum (s'il existe) de $A$ et par Max $A$ le plus grand élément ou élément maximum (s'il existe) de $A$. Les opérations Min et Max sont donc des opérations partielles (c'est-à-dire non partout définies) sur les parties de $X$ à valeurs dans $X$. Ces opérations peuvent être définies symboliquement:

$$\forall x\, [x \in A \Leftrightarrow \text{Min}\, A \leqslant x], \quad \forall x\, [x \in A \Leftrightarrow x \leqslant \text{Max}\, A].$$

Introduisons enfin les notations pour les bornes relativement à un ensemble $M \subset X$. On appelle *borne inférieure* (ou *intersection*) d'une partie $A$ d'un ensemble ordonné $\mathrm{X} = (X, \leqslant)$ relativement à un ensemble $M$ l'élément maximum (s'il existe) de l'ensemble $\beta_M^-(A)$ et on le note $\inf_M A$ ou encore $\bigcap^M A$, $\bigcap\limits_{x \in A}^M x$. Il est évident que $\inf_M A = \text{Max}\, \beta_M^-(A)$. On introduit de façon duale la notion de *borne supérieure* (ou *réunion*) d'un ensemble $A$ relativement à un ensemble $M$ et on la note $\sup_M A$ ou $\bigcup^M A$, $\bigcup\limits_{x \in A}^M x$. Ici $\sup_M A = = \text{Min}\ \beta_M^+(A)$.

Il est évident que les opérations Min et Max sont les restrictions des opérations inf et sup respectivement.

Voyons quelques types d'ensembles ordonnés qui nous serviront dans la suite. On appelle *inf* (resp. *sup*) *quasi-treillis* un ensemble ordonné dans lequel tout ensemble à deux éléments (ou ce qui revient au même, tout ensemble fini non vide) minoré (resp. majoré) admet une borne inférieure (resp. supérieure). On appelle *quasi-treillis* un ensemble ordonné qui est à la fois inf et sup quasi-treillis.

On appelle *inf* (resp. *sup*) *demi-treillis* un ensemble ordonné dans lequel tout ensemble à deux éléments (ou ce qui revient au même. tout ensemble fini non vide) admet une borne inférieure (resp. supérieure). On appelle *treillis* un ensemble ordonné qui est à la fois inf et sup demi-treillis.

Il est immédiat de voir qu'un ensemble ordonné est un inf (resp. sup) demi-treillis si et seulement s'il est filtrant inférieurement (resp. supérieurement) et si de plus il est un inf (resp. sup) quasi-treillis.

On appelle *treillis quasi complet* un ensemble ordonné dans lequel tout ensemble minoré non vide admet une borne inférieure ou ce qui revient au même tout ensemble majoré non vide admet une borne supérieure. Montrons que ces deux définitions sont identiques. En effet, supposons que dans un ensemble ordonné $(X, \leqslant)$ tout ensemble minoré non vide admet une borne supérieure et soit $M \subset X$ un ensemble majoré non vide. Prouvons que $M$ admet une borne supé-

rieure. Considérons un élément $m_0$ quelconque de $M$ (ce choix est possible puisque $M$ est non vide). L'ensemble $\beta^+(M)$ des majorants est minoré (par exemple, par $m_0$), donc il admet une borne inférieure $\inf \beta^+ (M)$ qui sera borne inférieure de $M$, ce qu'on vérifie immédiatement.

DÉFINITION. *On appelle fermeture* *) *sur un ensemble ordonné* $\mathbf{X} = (X, \leqslant)$ *toute application idempotente, croissante, dilatante* $P$ *sur le support* $X$ *de* $\mathbf{X}$, *c'est-à-dire une opération* $P$ *sur* $X$ *telle que soient réalisés les axiomes de la fermeture*:

$(C.1)_P$: $\forall x\ [P(P(x)) = P(x)]$ (*idempotence*);
$(C.2)_P$: $\forall x,\ y\ [x \leqslant y \Rightarrow P(x) \leqslant P(y)]$ (*croissance*);
$(C.3)_P$: $\forall x,\ [x \leqslant P(x)]$.

On a vu plus haut que l'idempotence de $P$ signifie que $P$ applique $X = \mathscr{D}(P)$ sur l'ensemble de ses points fixes: $\mathscr{R}(P) = P[X] = \mathrm{Fix}(P)$. Si $P$ est une fermeture, l'élément $P(x)$ sera dit *P-fermeture* de $x \in X$. Les éléments fixes de $P$ sur $\mathbf{X}$, c'est-à-dire les éléments de $P[X]$ sont dits *éléments fermés* de $P$.

Signalons que l'axiome $(C.1)_P$ est équivalent (sous l'hypothèse de la réalisabilité de $(C.3)_P$ à la propriété suivante:

$$(C.1)'_P: \forall x\ [P(P(x)) \leqslant P(x)].$$

Voyons maintenant les notions duales de celles que nous venons d'introduire.

On appelle *intérieur* **) sur un ensemble ordonné $\mathbf{X}$ une opération $P$ qui est fermeture sur l'ensemble ordonné $\mathbf{X}^{-1} = (X, \leqslant^{-1}) = (X, \geqslant)$ muni de l'ordre inverse. Si $P$ est un intérieur sur $\mathbf{X}$, ses éléments fixes sont dits *éléments ouverts* de $P$, et $P(x)$, *P-intérieur* (ou *noyau ouvert*) de $x$.

Les axiomes (C.1), (C.2) et (C.3) qui servent ordinairement à définir la notion de fermeture sont malgré leur simplicité assez algébriques et recèlent le sens intuitif de cette notion.

Nous allons donc exhiber un autre système d'axiomes de fermeture plus efficace du point de vue de l'établissement d'une classe assez naturelle de propriétés de la fermeture ***).

Ce système qui est composé de l'axiome $(C.3)_P$ et de l'axiome

$$(C.4)_P: \quad \forall x,\ y\ [x \leqslant P(y) \Rightarrow P(x) \leqslant P(y)]$$

s'appelle *principal*.

La démonstration de l'équivalence des deux systèmes sera omise, car triviale.

---

*) Dans les applications, les fermetures seront encore appelées projecteurs en haut ou enveloppes.

**) On l'appelle parfois projecteur en bas.

***) L'explication du sens exact et de la démonstration de cette assertion implique l'usage de méthodes de logique et d'algèbre.

Donc, le fait que l'opération $P$ sur $X$ est une fermeture sur X s'exprime maintenant à l'aide de la condition $(C.3)_P$ & $(C.4)_P$, qui signifie de toute évidence que la fermeture $P(x)$ de tout élément $x \in X$ est le plus petit de tous les éléments $P$-fermés majorant $x$:

$$(C.5)_P: (\forall x \in X) [P(x) = \operatorname{Min} (P[X] \cap [x, \rightarrow)) = \operatorname{Min} (\operatorname{Fix}(P) \cap [x, \rightarrow))].$$

Il est utile d'envisager la relation suivante entre les opérations $P$ sur $X$ et les ensembles $M \subset X$:

$$(C.6)_{P, M}: (\forall x \in X) [P(x) = \operatorname{Min} (M \cap [x, \rightarrow))].$$

Si pour un ensemble $M \subset X$ il existe une opération $P$ sur $X$ liée à l'ensemble $M$ par la relation $(C.6)_{P, M}$, alors cette opération est visiblement définie de façon unique par $M$. Ceci étant, on vérifie immédiatement que $P$ est une fermeture sur l'ensemble ordonné $\mathbf{X} = (X, \leqslant)$, et $M$ le système de tous les éléments fermés de $P$: $M = P[X] = \operatorname{Fix}(P)$. Nous ne donnons pas la démonstration de ce fait simple, car c'est un cas particulier d'une proposition plus générale qui sera prouvée au n° 1.4 (théorème 1).

De ce qui précède il suit le

THÉORÈME 1. *Pour que dans un ensemble ordonné* $\mathbf{X} = (X, \leqslant)$ *un ensemble* $M$ *soit représentable par un système d'éléments fermés d'une opération de fermeture* $P$ *sur* $\mathbf{X}$, *il est nécessaire et suffisant que sur chaque coupe inférieure* $M_x = M \cap [x, \rightarrow)$ *de cet ensemble* $M$ *suivant* $x \in X$ *il existe un plus petit élément, c'est-à-dire que l'on ait la condition*

$$(C.7)_M: (\forall x \in X) (\exists y \in X) [y = \operatorname{Min} (M \cap [x, \rightarrow))].$$

*Si en outre cette condition est réalisée sur* $M$, *alors la fermeture* $P$, *qui est mentionnée dans le théorème, est définie de façon unique par l'ensemble* $M: P(x) = \operatorname{Min} (M \cap [x, \rightarrow))$, $\forall x \in X$.

◀ Pour démontrer ce théorème (compte tenu de ce qui a été dit avant) il reste à remarquer que la condition $(C.7)_M$ sur $M$ équivaut à l'existence d'une opération $P$ sur $X$ liée à cet ensemble par la relation $(C.6)_{P, M}$, c'est-à-dire équivaut à la condition:

$$(\exists P: X \rightarrow X) (\forall x \in X) [P(x) = \operatorname{Min} (M \cap [x, \rightarrow))].$$

ce qui est intuitif *). ▶

Signalons que la conjonction $(C.3)_P$ & $(C.4)_P$ équivaut manifestement à la condition suivante sur l'opération $P$:

$$(C.8)_P: \forall x, y [x \leqslant P(y) \Leftrightarrow P(x) \leqslant P(y)].$$

*) On peut du reste le prouver d'une façon formelle en remarquant que d'une manière générale une proposition de la forme: $(\forall x \in A) (\exists y \in B) R(x, y)$, où $R(x, y)$ est une proposition « dépendant » des variables $x$ et $y$, est équivalente à: $(\exists f: A \rightarrow B) (\forall x \in A) R(x, f(x))$.

Dans un langage moins formel, $P$ est une fermeture sur X si et seulement si tout élément $x \in X$ minore les $P$-images, et elles seules, qui sont minorées par la fermeture $P(x)$ de $x$; il est équivalent de dire que:

$$(\forall A \subset P[X]) \quad [P[\beta_X^-(A) = \beta_{P[X]}^-(A)].$$

Il résulte de la condition $(C.8)_P$ que les opérations $\inf_X$ et $\inf_{P[X]}$ sont confondues sur les parties de $P[X]$:

$$(\forall A \subset P[X]) \quad [\inf_X(A) \rightleftarrows \inf_{P[X]}(A)].$$

A propos du signe $\rightleftarrows$ (souvent désigné par $\simeq$) on adoptera la convention suivante. Si $T$ et $U$ sont deux termes quelconques, l'égalité $T \rightleftarrows U$ signifie que si la valeur d'un membre de cette égalité est déterminée, celle de l'autre le sera également et ces deux valeurs sont confondues. Cette notation est d'une grande utilité dans les raisonnements sur les opérations partielles. Nous nous proposons à cet effet de la définir avec plus de soin. Nous définissons $T \rightleftarrows U$ comme la relation: $\forall x\, [T = x \Leftrightarrow U = x]$. où les termes $T$ et $U$ ne dépendent pas de $x$. Nous aurons aussi besoin de l'égalité $T \rightrightarrows U$ qui se définit par la relation: $\forall x\, [T = x \Rightarrow U = x]$. Généralement, l'égalité forte $T = U$ est comprise comme suit: $\exists x\, [T = x = U)$.

Nous nous proposons de donner une définition qui nous permettra de formuler un critère utile affirmant qu'un ensemble $M \subset X$ est l'ensemble des éléments fermés d'une fermeture sur $(X, \leqslant)$.

DÉFINITION. *On dit qu'un ensemble $M$ d'un ensemble ordonné* $\mathbf{X} = (X, \leqslant)$ *est régulièrement plongé (vers le bas) dans* $\mathbf{X}$ *si pour toute partie $A$ de $M$ la trace $\beta^-(A) \cap M$ est cofinale dans $\beta^-(A)$, c'est-à-dire pour tout minorant $x$ de $A$ on peut exhiber un minorant $y$ de $A$ appartenant à $M$ tel que $x \leqslant y$.*

La régularité de l'immersion vers le bas de l'ensemble $M$ dans $\mathbf{X}$ signifie que toute partie $A$ de $M$ est approchée inférieurement par les éléments de $M$ pas plus mal que par les éléments de $X$.

THÉORÈME 2. *Pour qu'il existe une fermeture sur un ensemble ordonné* $\mathbf{X} = (X, \leqslant)$ *dont l'ensemble des éléments fermés est $M$, il est nécessaire et suffisant que $M$ soit régulièrement plongé (vers le bas) dans* $\mathbf{X}$.

◀ D'après les propriétés de la fermeture, seule la condition suffisante est à démontrer. Supposons que $M$ est régulièrement plongé (vers le bas) dans $\mathbf{X} = (X, \leqslant)$. En vertu du théorème 1, il nous suffit seulement de prouver que pour tout $x \in X$ la coupe inférieure $M_x$ admet un plus petit élément. Soit $x \in X$. Il est évident que $x$ est un minorant pour l'ensemble $M_x$. L'ensemble $M$ étant régulièrement plongé (vers le bas) dans $\mathbf{X}$ il doit contenir un élément $y \geqslant x$ qui est aussi un minorant pour $M_x$. Comme $y \in M$ et $y \geqslant x$, il vient $y \in$ $\in M_x = M \cap [x, \rightarrow)$ et le minorant $y$ sera élément minimum de $M_x$. ▶

Voyons maintenant comment se transforment les bornes par une application de X dans X, qui est une fermeture.

On remarquera pour commencer que si l'opération $P$ sur X est croissante, alors

$$P[\beta^+(A)] \subset \beta^+(P[A]) \cap P[X] = \beta^+_{P[X]} P[A],$$

$$P[\beta^-(A)] \subset \beta^-(P[A]) \cap P[X] = B^-_{P[X]} P[A].$$

THÉORÈME 3. *Si $P$ est une fermeture sur un ensemble ordonné* $X = (X, \leqslant)$, *alors*

(1) $P[\beta^-_X(A)] = \beta^-_{P[X]}(A), \ \forall A \subset P[X]$ ;

(2) $\inf_X(A) \rightleftarrows \inf_{P[X]}(A), \ \forall A \subset P[X]$ ;

(3) $P[\beta^+_X(A)] = \beta^+_{P[X]}(P[A]), \ \forall A \subset X$ ;

(4) $P(\sup_X A) \rightleftarrows \sup_{P[X]} P[A], \ \forall A \subset X$ ;

(5) $\sup_{P[X]}(A) \rightleftarrows \inf_X P[\beta^+_X(A)], \ \forall A \subset P[X]$.

◀ Comme indiqué précédemment, les relations (1) et (2) sont une conséquence immédiate de la condition $(C.8)_P$ sur l'opération $P: \forall x, y \ [Px \leqslant Py \Leftrightarrow x \leqslant Py]$.

La condition (3) est remplie, puisqu'elle est l'écriture symbolique du fait simple que tout élément de $X$ est fermeture d'un majorant de $A$ si et seulement s'il est majorant fermé de $A$.

Prouvons maintenant la condition (4). Soient $A \subset X$, $y = \sup A$. Montrons que $P(y) = \sup_{P[X]} P[A]$. L'application $P$ étant croissante et $y = \sup A$ il s'ensuit que $P(y)$ est un majorant fermé pour $P[A]$. Supposons maintenant que $z$ est un majorant fermé quelconque pour $P[A]$, c'est-à-dire $(\forall a \in A)\ [a \leqslant z = P(z)]$. Montrons que $P(y) \leqslant z$. (En effet, puisque $y = \sup A$ et $z \in \beta^+ A$, on a $y \leqslant z$. D'où, en vertu de la fermeture de $z$ et de la propriété (C.4), il suit: $P(y) \leqslant z$.

Nous allons ébaucher les grandes lignes de la démonstration de la condition (5) et laisser au lecteur le soin des détails. Si $A \subset P[X]$, alors $P[A] = A$. D'autre part $\sup_{P[X]} A$ est confondu par définition avec l'élément minimum de l'ensemble $\beta^+_{P[X]} A : \sup_{P[X]} A \rightleftarrows \rightleftarrows \mathrm{Min}\, \beta^+_{P[X]} A$. Il faut ensuite se servir des relations (1) à (4) déjà prouvées. ▶

Ce théorème appelle quelques commentaires. Le numéro 2 dit que les opérations $\inf_X$ et $\inf_{P[X]}$ sont confondues sur les ensembles des éléments fermés (corollaire de la condition (C.8)). Dans le cas particulier où $A \subset P[X]$, (4) implique:

$$(6) \quad \sup_{P[X]} A \leftleftarrows P(\sup_X A), \ \forall A \subset P[X],$$

ce qui signifie que l'opération $\sup_{P[X]}$ est le prolongement de l'opération de prise de la borne supérieure $P$-fermée : $P \circ \sup_X$.

Si $\sup_{F[X]}$ est définie et $P(\sup_X A)$ ne l'est pas (il existe des exemples simples d'une telle situation), la condition (4) ne nous fournit aucune information sur $\sup_{F[X]} A$. Il n'empêche que l'assertion (5) qui donne l'expression de $\sup_{F[X]} A$ valable pour tous les cas, mais, il est vrai, en termes des opérations $\inf_X$ et $\beta_X^+$, présente de l'intérêt. Signalons que la condition (2) entraîne immédiatement le

COROLLAIRE. *L'ensemble des éléments fermés d'une fermeture quelconque sur un ensemble ordonné est stable pour les opérations (partielles) inf (on dit encore stable pour les intersections), c'est-à-dire la borne inférieure (resp. l'intersection) de tout ensemble d'éléments fermés, si elle existe, est aussi un élément fermé.*

A noter que la stabilité d'une partie d'un ensemble ordonné pour les opérations inf ne suffit pas à elle seule à faire de cette partie l'ensemble des éléments fermés d'une fermeture. Il y a de ce fait intérêt à reformuler le critère (le théorème 1) de représentabilité d'une partie d'un ensemble ordonné par l'ensemble des éléments fermés d'une fermeture en termes de stabilité pour l'intersection.

THÉORÈME 1 bis. *Pour qu'une partie $M$ d'un ensemble ordonné $\mathbf{X} = (X, \leqslant)$ soit représentable par l'ensemble des éléments fermés d'une fermeture sur $\mathbf{X}$, il est nécessaire et suffisant que la borne inférieure de toute coupe inférieure de $M$ soit définie et appartienne à $M$, c'est-à-dire* $(\forall x \in X)$ $[\inf\ (M \cap [x, \rightarrow)) \in M]$.

REMARQUE. Si l'ensemble ordonné $\mathbf{X}$ est un treillis complet, alors un ensemble $M \subset X$ se représente par l'ensemble des éléments fermés d'une fermeture sur $\mathbf{X}$ si et seulement si cet ensemble est stable pour l'opération inf. Ceci est une conséquence immédiate de ce qui a été dit plus haut.

On est souvent amené à considérer plusieurs fermetures sur un même ensemble ordonné (en général un ensemble ordonné par l'inclusion de la forme $(\mathfrak{P}(S), \subset)$, où $S$ est un espace, et à établir sous quelles conditions une fermeture est la composée de deux autres.

THÉORÈME 4. *Soient $\Pi_1$ et $\Pi_2$ deux fermetures sur un ensemble ordonné $\mathbf{X} = (X, \leqslant)$. Les conditions suivantes sont équivalentes :*

1) *La composée $\Pi_1 \circ \Pi_2$ est une fermeture sur $\mathbf{X}$ ;*

2) *L'ensemble des éléments fermés de $\Pi_2$ est stable pour la fermeture $\Pi_1$, autrement dit $\Pi_1[\Pi_2[X]] \subset \Pi_2[X]$ ;*

3) $\Pi_2 \circ \Pi_1 \circ \Pi_2 = \Pi_1 \circ \Pi_2$.

4) $(\Pi_1 \circ \Pi_2)[X] = \Pi_1[X] \cap \Pi_2[X]$.

DÉMONSTRATION. Les conditions 2) et 3) sont équivalentes, puisque l'inclusion $A \subset \Pi_2[X]$ signifie que tout élément de $A$ est un point fixe pour la fermeture $\Pi_2$. Prouvons maintenant que 3) entraîne 1). L'application $\Pi_1 \circ \Pi_2$ est croissante et dilatante, puisque $\Pi_1$ et

$\Pi_2$ le sont. Pour établir l'idempotence de $\Pi_1 \circ \Pi_2$ il suffit de remarquer que $\Pi_1 \circ \Pi_2 \circ \Pi_1 \circ \Pi_2 = \Pi_1 \circ \Pi_1 \circ \Pi_2 = \Pi_1 \circ \Pi_2$.

Montrons enfin que 1) entraîne 3). Remarquons tout d'abord que pour tous les $x \in X$ on a $(\Pi_1 \circ \Pi_2)(x) \leqslant (\Pi_2 \circ \Pi_1 \circ \Pi_2)(x)$. Prouvons que la relation inverse est vraie. En effet, $(\Pi_1 \circ \Pi_2)(x) = (\Pi_1 \circ \Pi_2 \circ \Pi_1 \circ \Pi_2)(x) \geqslant (\Pi_2 \circ \Pi_1 \circ \Pi_2)(x)$. On obtient donc

$$(\forall x \in X)\,[(\Pi_2 \circ \Pi_1 \circ \Pi_2)(x) = (\Pi_1 \circ \Pi_2)(x)],$$

c'est-à-dire la condition 3).

Pour achever la démonstration il reste à remarquer que l'égalité

$$(\Pi_1 \circ \Pi_2)[X] = \Pi_1[X] \cap \Pi_2[X]$$

est une conséquence évidente des conditions 1) et 3), et inversement, la condition 2) est une conséquence de 4). ▶

A noter que même si la composée $\Pi_1 \circ \Pi_2$ des fermetures $\Pi_1$ et $\Pi_2$ est une fermeture, la composée $\Pi_2 \circ \Pi_1$ n'est pas en général une fermeture et donc n'est pas confondue avec $\Pi_1 \circ \Pi_2$. Plus bas nous aurons affaire à des exemples de cette nature. La permutabilité de deux fermetures entraîne la réalisation de la condition 1), donc des conditions 2), 3) du théorème 4, puisque

$$\Pi_1[\Pi_2[X]] = \Pi_2[\Pi_1[X]] \subset \Pi_2[X]$$

On a vu au début que la notion duale de la fermeture sur un ensemble ordonné $(X, \leqslant) = \mathrm{X}$ est la notion d'intérieur sur cet ensemble, plus exactement, une opération $P$ sur $X$ est un intérieur sur $\mathrm{X}$ si est réalisé le système suivant d'axiomes de l'intérieur :

(1.1) $\forall x\,[P(P(x)) = P(x)]$ (idempotence),

(1.2) $\forall x, y\,[x \leqslant y \Rightarrow P(x) \leqslant P(y)]$ (croissance),

(1.3) $\forall x\,[P(x) \leqslant x]$.

Il est bien entendu que la théorie de la fermeture sur un ensemble ordonné peut être considérée comme la théorie de l'intérieur sur un ensemble ordonné si l'on remplace dans la première l'ordre $\leqslant$ qui est arbitraire par l'ordre inverse $\geqslant$. Donc ces deux théories sont isomorphes, l'ordre dual de $\leqslant$ étant $\geqslant$ et les opérations duales de $\beta^-$, $\beta^+$, inf, sup respectivement $\beta^+$, $\beta^-$, sup, inf. C'est la raison pour laquelle nous n'avons pas donné les formulations duales.

Voyons maintenant le cas, le plus important pour nous, où pour ensemble ordonné on prend l'ensemble des parties $\mathfrak{P}(S)$ d'un ensemble $S$ ordonné par l'inclusion. Il est évident que cet ensemble ordonné est un treillis complet dont les bornes inférieure et supérieure sont confondues avec l'intersection et la réunion des ensembles :

$$\inf \mathfrak{A} = \bigcap \mathfrak{A} = \bigcap_{A \in \mathfrak{A}} A, \quad \sup \mathfrak{A} = \bigcap \mathfrak{A} = \bigcup_{A \in \mathfrak{A}} A,$$

où $\mathfrak{A}$ est un ensemble quelconque de parties de $S$, c'est-à-dire $\mathfrak{A} \subset \mathfrak{P}(S)$. Les fermetures sur l'ensemble ordonné $(\mathfrak{P}(S), \subset)$ seront encore appelées *opérateurs de fermeture* sur $S$. D'après ce qui précède

les opérateurs de fermeture $P$ sur $S$ sont en bijection avec les systèmes de parties de $S$, stables pour les intersections (ces systèmes seront appelés brièvement *systèmes* $\bigcap$ stables ou $\bigcap$-systèmes):

$$(\forall \mathfrak{B} \subset \mathfrak{A}) [\bigcap \mathfrak{B} = \bigcap_{B \in \mathfrak{B}} B \in \mathfrak{A}].$$

De plus à tout opérateur de fermeture $P$ sur $S$ est associé l'ensemble des parties $P$-fermées de $S$, c'est-à-dire l'ensemble $\{A \subset S : P(A) = A\} = \text{Fix}(P)$. Inversement, à tout ensemble $\mathfrak{A}$ de parties de $S$ stable pour l'intersection est associé un opérateur de fermeture $A \mapsto \bigcap \bigcap \{B \in \mathfrak{A} : A \subset B\}$ sur $S$ que l'on notera $\mathscr{C}_{\mathfrak{A}}$. On dira aussi que $\mathscr{C}_{\mathfrak{A}}(A)$ est $\mathfrak{A}$-enveloppe de $A$. Si le système $\mathfrak{A}$ est $\bigcap$-stable, celle-ci sera évidemment égale au plus petit des ensembles de $\mathfrak{A}$ qui contiennent $A$. Comme l'opération de complémentation des parties de $S$ (que nous désignerons par $\nu$) est une involution réalisant un isomorphisme inverse de l'ensemble ordonné $\mathfrak{X} = (\mathfrak{P}(S), \subset)$ dans lui-même, c'est-à-dire $\nu(\nu(A)) = A$,

$$A \subset B \Leftrightarrow \nu(A) \subset \nu(B), \quad \text{pour tous } A, B \subset S,$$

il existe une application bijective qui à toute fermeture $P$ sur $\mathfrak{X}$ associe l'intérieur dual $I = \nu \circ P \circ \nu$ et inversement à tout intérieur $I$ sur $\mathfrak{X}$ associe la fermeture $P = \nu \circ I \circ \nu$. Cette application associe également à tout système $\bigcap$-stable de parties de $\mathfrak{A}$ un système $\bigcup$-stable de parties de $\nu[\mathfrak{A}] = \{\nu(A) : A \in \mathfrak{A}\}$ et inversement.

Exhibons un exemple simple d'opérateur de fermeture sur un ensemble $S$. Soit $f$ une application de $S$ sur un ensemble $Q$. Désignons par $P_f$ l'opération $A \mapsto f^{-1}[f[A]]$, $A \subset S$, qui applique l'ensemble des parties $\mathfrak{P}(S)$ dans lui-même. L'ensemble $P_f(A) = f^{-1}[f[A]]$ s'appelle *saturé* de l'ensemble $A$ pour l'application $f$ et l'opération $P_f$ *opération de saturation* pour l'application $f$. Montrons que cette opération est un opérateur de fermeture sur l'ensemble $S$. On conviendra de désigner par $\hat{F}$ le passage à l'image par une correspondance $F$. D'après le n° 1.1 on a alors $f \circ f^{-1} = I_Q$, donc $\hat{f} \circ \widehat{f^{-1}} = I_{\mathfrak{P}(Q)}$. Par suite, l'application $\hat{f}$ est l'inverse à gauche de l'application $\widehat{f^{-1}}$ qui en vertu du n° 1.1 est injective. La composée $\widehat{f^{-1}} \circ \hat{f} = P_f$ sera d'après le n° 1.1 un projecteur (une application idempotente) de l'ensemble $\mathfrak{P}(S)$ sur l'ensemble des parties de $S$ représentables par $f^{-1}[B]$, où $B \subset Q$. L'application $P_f$ est en outre croissante et dilatante puisqu'elle est la composée des applications $\widehat{f^{-1}}$, $\hat{f}$ qui le sont. On a donc prouvé que $P_f$ est un opérateur de fermeture sur $S$. Si l'on introduit la relation canonique d'équivalence $\overset{\approx}{f}$ sur l'ensemble $S$ en convenant que $x \overset{\approx}{f} y \Leftrightarrow (x) = f(y)$ pour tous les $x, y \in S$, on obtient que l'ensemble $A \subset S$ est fermé: $P_f(A) = A$ ou comme on dit en pareil cas l'ensemble $A$ est saturé pour l'application $f$ si et seule-

ment s'il est compatible avec la relation d'équivalence canonique $\tilde{\tilde{f}}$ c'est-à-dire $(x \in A)$ & $(x \tilde{\tilde{f}} y) \Rightarrow (y \in A)$ pour tous $x, y \in S$. Si de plus l'on désigne par $[x]_f$ la classe d'équivalence d'un élément $x$ pour la relation d'équivalence $\tilde{\tilde{f}}$, c'est-à-dire $[x]_f = \{y \in S : x \tilde{\tilde{f}} y\}$, il vient alors que le saturé de tout ensemble est tout simplement la réunion des classes d'équivalence de tous les éléments de $S$ pour la relation d'équivalence $\tilde{\tilde{f}}$, c'est-à-dire

$$P_f(A) = \bigcup_{x \in A} [x]_f = \bigcup_{x \in A} P_f\{x\} \quad \text{pour tous les} \quad A \subset S.$$

Il est clair que l'ensemble de toutes les parties de $S$ saturées pour $f$ sera stable pour l'intersection de même que pour la réunion.

**1.3.** Dans ce numéro on introduit deux types d'opérateurs de fermeture sur un ensemble: un opérateur topologique et un opérateur algébrique. On exhibe certains faits classiques d'algèbre sur la caractérisation de systèmes d'ensembles fermés pour de tels opérateurs de fermeture de même qu'on établit une relation entre les opérateurs de fermeture algébriques et les algèbres universelles.

PROPOSITION 1. *Entre les opérateurs de fermeture $P$ sur un ensemble $S$ préservant les réunions finies $(TP)_P$: $P(\bigcup_{A \in \mathfrak{A}} A) = \bigcup_{A \in \mathfrak{A}} P(A)$ pour tous les $\mathfrak{A} \subset \mathfrak{P}(S)$ finis, et les systèmes $\mathfrak{B}$ de parties de l'ensemble $S$ stables pour des intersections quelconques et des réunions finies il existe une correspondance biunivoque définie par le couple de fonctions: $P \mapsto$ Fix $(P)$ et $\mathfrak{B} \mapsto \mathscr{C}_{\mathfrak{B}}$.*

DEMONSTRATION. En effet, tout système fini de « points» fixes» de l'opérateur de fermeture $P$ peut être mis sous la forme: $\{P(A) : A \in \in \mathfrak{A}\}$, où $\mathfrak{A}$ est un système fini de parties de l'ensemble $S$. Par suite de la propriété $(TP)_P$ on a: $\bigcup \{P(A) : A \in \mathfrak{A}\} = P(\bigcup \{A : A \in \mathfrak{A}\}) = P(\bigcup \mathfrak{A}) \in \text{Fix}(P)$, ce qui exprime que le système Fix $(P)$ est stable pour les réunions finies. La $\bigcap$-stabilité de Fix $(P)$ a été établie antérieurement dans le cas général sans l'hypothèse $(TP)_P$. Inversement, si $\mathfrak{B}$ est un système stable pour des intersections quelconques et des réunions finies, alors pour tout sous-système fini $\mathfrak{A} \subset \mathfrak{B}$ on a également $\bigcup \mathfrak{A} \in \mathfrak{B}$, donc $\mathscr{C}_{\mathfrak{B}}(\bigcup \mathfrak{A}) = \bigcup \mathfrak{A} = \bigcup \{\mathscr{C}_{\mathfrak{B}}(A) : A \in \mathfrak{A}\}$, puisque Fix $(\mathscr{C}_{\mathfrak{B}}) = \mathfrak{A}$. Par suite $\mathscr{C}_{\mathfrak{B}}$ préserve les réunions finies. ▶

Les opérateurs de fermeture sur un ensemble $S$ préservant les réunions finies sont dits *opérateurs de fermeture topologiques* sur $S$, et les systèmes de parties fermées de $S$ qui leur sont associés, c'est-à-dire les systèmes de parties stables pour des intersections quelconques et des réunions finies, *systèmes topologiques de parties fermées* ou *topologies* sur $S$ (sous forme de système de parties fermées). La condition $(TP)_P$ sur la fermeture $P$ s'écrit généralement sous la forme d'une conjonction de deux conditions:

(K.0) $P\,\varnothing = \varnothing$,
(K.2) $P\,(A_1 \cup A_2) = P\,(A_1) \cup P\,(A_2)$, $\forall A_1,\ A_2 \subset S$.
Ces deux conditions jointes aux axiomes de fermeture de Moore pour $P$:
(C.1) $P\,(P\,(A)) = P\,(A)$, $\forall A \subset S$,
(C.2) $A \subset B \Rightarrow P\,(A) \subset P\,(B)$, $\forall A,\ B \subset S$.
(C.3) $A \subset P\,(A)$, $\forall A \subset S$.
donnent le système d'axiomes de Kuratowski qui caractérisent la notion de fermeture topologique. Signalons qu'on obtiendrait un système d'axiomes équivalents en remplaçant le signe = par $\subset$ dans les axiomes (K.0), (K.2), (C.1).

Un opérateur de fermeture $P$ sur un ensemble $S$ est *algébrique* si pour tous $A \subset S$ et $x \in S$ l'appartenance $x \in P\,(A)$ entraîne l'existence d'une partie finie $F$ de l'ensemble $A$ telle que $x \in P\,(F)$, c'est-à-dire pour tout $A \subset S$:

$$P\,(A) = \bigcup \{P\,(F) : (F \subset A)\ \&\ (F \text{ est finie})\}.$$

Pour caractériser les systèmes de parties fermées pour des opérateurs de fermeture algébriques nous aurons besoin de la notion de système inductif de parties, ainsi que de la notion de système de caractère fini. Mais voyons tout d'abord le fait suivant basé sur le lemme de Zorn.

PROPOSITION. 2. *Soit $\mathfrak{B}$ un système quelconque de parties d'un ensemble $S$. Les trois conditions suivantes sont équivalentes*:

(1) *Le système $\mathfrak{B}$ est stable pour les réunions de sous-systèmes quelconques filtrant supérieurement (pour l'inclusion $\subset$) du système $\mathfrak{B}$*;

(2) *Le système $\mathfrak{B}$ est stable pour les réunions de sous-systèmes quelconques linéairement ordonnés (par l'inclusion $\subset$) de $\mathfrak{B}$*;

(3) *Le système $\mathfrak{B}$ est stable pour les réunions de sous-systèmes quelconques complètement ordonnés (par l'inclusion $\subset$) de $\mathfrak{B}$.*

On dit qu'un système $\mathfrak{B}$ de parties d'un ensemble $S$ est *inductif à droite* s'il vérifie la condition (1) (donc les conditions (2) et (3)). Ainsi donc, un système d'ensembles est inductif si chacun de ses sous-systèmes filtrant supérieurement est de nouveau un ensemble de ce système. On a le

THÉORÈME 5. *Un opérateur de fermeture sur un ensemble $S$ est algébrique si et seulement si le système de parties fermées de $S$ qui lui correspond est inductif.*

DÉMONSTRATION. Soient $P$ un opérateur algébrique sur $S$, $\mathfrak{A}$ un système filtrant supérieurement de parties fermées pour $P$: $\mathfrak{A} \subset$ $\subset$ Fix $(P)$. Prouvons que la réunion $\bigcup \mathfrak{A}$ sera également fermée pour $P$, c'est-à-dire $P\,(\bigcup \mathfrak{A}) \subset (\bigcup \mathfrak{A})$. En effet, soit $x \in P\,(\bigcup \mathfrak{A})$. L'opérateur $P$ étant algébrique, on peut trouver un sous-ensemble

fini $F = \{a_1, \ldots, a_n\}$ de $\bigcup \mathfrak{A}$ tel que $x \in P(F)$. Vu que $F \subset \bigcup \mathfrak{A}$ on obtient $a_i \in \bigcup \mathfrak{A}$, $\forall i = 1, \ldots, n$. D'où l'on déduit qu'il existe des ensembles $A_1, \ldots, A_n$ du système $\mathfrak{A}$ tels que $a_1 \in A_1, \ldots, a_n \in A_n$. Comme $\mathfrak{A}$ est filtrant supérieurement il existe dans le système $\mathfrak{A}$ un ensemble $B$ contenant chacun des ensembles $A_1, \ldots, A_n$. Il est évident que $F \subset A_1 \cup \ldots \cup A_n \subset B \subset \bigcup \mathfrak{A}$. D'où $P(F) \subset$ $\subset P(B) = B \subset \bigcup \mathfrak{A}$, puisque $P$ est croissant. Donc $x \in \bigcup \mathfrak{A}$.

Inversement, supposons maintenant que le système Fix $(P)$ est inductif. Montrons que pour tout $A \subset S$ on aura: $P(A) =$ $= \bigcup \{P(F) : (F \subset A)\ \&\ (F \text{ est finie})\}$. Désignons par $\mathfrak{B}_A$ le système d'ensembles $\{P(F) : (F \subset A) \& (F \text{ est finie})\}$. Prouvons que $\bigcup \mathfrak{B}_A \in$ Fix $(P)$, c'est-à-dire $\bigcup \mathfrak{B}_A$ est un ensemble fermé pour $P$. On remarquera que pour tous $B_1$, $B_2 \subset S$ on a: $P(B_1) \cup P(B_2) \subset$ $\subset P(B_1 \cup B_2)$, et si $B_1$, $B_2$ sont des parties finies de $A$, alors $B_1 \cup$ $\cup B_2$ est aussi une partie finie de $A$. Donc, le système $\mathfrak{B}_A$ est filtrant supérieurement et par suite sa réunion est fermée : $P(\bigcup \mathfrak{B}_A) =$ $= \bigcup \mathfrak{B}_A$. Puisque $A \subset \bigcup \mathfrak{B}_A$, on obtient en définitive $P(A) \subset$ $\subset P(\bigcup \mathfrak{A}_A) = \bigcup \mathfrak{A}_A$. L'inclusion inverse $\bigcup \mathfrak{A}_A \subset P(A)$ résulte de la croissance de $P$. On obtient en définitive $P(A) = \bigcup \mathfrak{A}_A$ pour tout $A \subset S$, ce qui exprime que $P$ est algébrique.

Un système $\mathfrak{A}$ de parties d'un ensemble $S$ s'appelle *système de caractère fini* s'il existe un système $\mathcal{F}$ de parties finies de $S$ tel que tout ensemble $B \in \mathfrak{A}$ est défini de façon unique par ses intersections $B \cap F$, où l'ensemble $F$ parcourt le système $\mathcal{F}$.

On démontre qu'un opérateur de fermeture sur un ensemble est algébrique si et seulement si le système de parties fermées qui lui correspond est un système de caractère fini.

Introduisons maintenant la notion de $\Omega$-algèbre ou algèbre (universelle) de signature $\Omega$. Soit donné un ensemble $\Phi$ dont les éléments seront appelés symboles d'opérations et supposons qu'à chaque symbole $\varphi \in \Phi$ est associé (par une application fixée $\tau$) un entier naturel $\tau(\varphi) \geqslant 0$ appelé nombre d'arguments de l'opération $\varphi$. Appelons *signature* toute partie $\Omega$ d'un ensemble de symboles de $\Phi$. Supposons que $\Omega$ est une signature arbitraire ; c'est-à-dire $\Omega \subset \Phi$. Par *algèbre de signature* $\Omega$ (ou *$\Omega$-algèbre*) on entend tout couple de la forme $\mathfrak{A} = (S, \mathcal{A})$, où $S$ est un ensemble non vide appelé *support* de l'algèbre $\mathfrak{A}$ et $\mathcal{A}$ une fonction définie sur la signature $\Omega$ et associant à chaque symbole de $\varphi \in \Omega$ une opération $\tau(\varphi)$-aire $\mathcal{A}(\varphi)$ sur le support $A$, c'est-à-dire $\mathcal{A}(\varphi) : A^{\tau(\varphi)} \to A$, $\forall \varphi \in \Omega$. En outre, au lieu de $\mathcal{A}(\varphi)$ on écrit aussi : $\varphi^{\mathfrak{A}}$ ou $\varphi_{\mathfrak{A}}$ et on dit que $\mathcal{A}(\varphi) = \varphi^{\mathfrak{A}} = \varphi_{\mathfrak{A}}$ est la réalisation (ou l'estimation) du symbole de l'opération $\varphi$ dans $\mathfrak{A}$. Toute partie $A$ de $S$ stable pour les opérations $\varphi^{\mathfrak{A}}$, $\varphi \in \Omega$, peut être considérée comme une $\Omega$-algèbre si à chaque $\varphi \in \Omega$ on associe la restriction de l'opération $\varphi^{\mathfrak{A}}$ à cette partie. Cette $\Omega$-algèbre sur le support $A$ s'appelle *$\Omega$-sous-algèbre* de $\mathfrak{A}$.

Le système de toutes les $\Omega$-sous-algèbres de $\mathfrak{A}$, qui sera désigné dans la suite par $\mathscr{B}_\Omega(\mathfrak{A})$, est stable pour les intersections. Nous désignerons par $J_\mathfrak{A}$ l'opérateur de fermeture défini par ce système sur le support de l'algèbre $\mathfrak{A}$. Si $A \subset S$, alors $J_\mathfrak{A}(A)$ est la $\Omega$-sous-algèbre de l'algèbre $\mathfrak{A}$ engendrée par l'ensemble $A$, c'est-à-dire la plus petite de toutes les sous-algèbres de $\mathfrak{A}$ contenant $A$. Nous aurons besoin des notions suivantes pour décrire la structure de la sous-algèbre $J_\mathfrak{A}(A)$. On appelle *terme* d'une signature $\Omega$ des expressions formelles obtenues à partir de variables formelles $x_1, x_2, \ldots$ données *a priori* par une application multiple des symboles des opérations de $\Omega$ *). On dit que $\theta$ est l'estimation (des variables) du terme $T$ dans l'algèbre $\mathfrak{A}$ si $\theta$ est une fonction associant à toute variable de $T$ un élément (le support) de l'algèbre $\mathfrak{A}$. Si enfin $\theta$ est une estimation du terme $T$ dans l'algèbre $\mathfrak{A}$, on désignera alors par $v_{\mathfrak{A},\theta}(T)$ la valeur de ce terme dans l'algèbre $\mathfrak{A}$ lorsqu'on estime ses variables $\theta$, c'est-à-dire le résultat du «calcul» de $T$ relatif à l'interprétation des symboles des opérations et des variables formelles de $T$, donnée par l'algèbre $\mathfrak{A}$ et l'estimation $\theta$.

PROPOSITION 3. *Soit* $\mathfrak{A} = (S, \mathscr{A})$ *une* $\Omega$*-algèbre sur* $S$. *Pour tout* $A \subset S$, *la sous-algèbre de* $\mathfrak{A}$ *engendrée par* $A$ *est confondue avec l'ensemble de valeurs des termes de la signature* $\Omega$ *dans l'algèbre* $\mathfrak{A}$ *par toutes les estimations des variables de ces termes par les éléments du sous-ensemble* $A$ *du support de l'algèbre* $\mathfrak{A}$ : $J_\mathfrak{A}(A) = \{v_{\mathfrak{A},\theta}(T) : T$ *est le terme de signature* $\Omega$, $\theta$ *l'estimation de* $T$ *dans le sous-ensemble* $A$ *de l'algèbre* $\mathfrak{A}\}$.

L'opérateur de fermeture $J_\mathfrak{A}$ sera visiblement algébrique, puisque chaque terme ne contient qu'un ensemble fini de variables. Cette propositon et sa réciproque constituent le théorème suivant prouvé par P. Hall et J. Schmidt.

THÉORÈME 6. *Pour qu'un opérateur de fermeture* $P$ *sur un ensemble* $S$ *soit algébrique il est nécessaire et suffisant qu'il soit confondu avec l'opérateur* $J_\mathfrak{A}$ *de génération des sous-algèbres d'une algèbre universelle* $\mathfrak{A}$ *sur* $S$ : $P(A) = J_\mathfrak{A}(A)$ *est la sous-algèbre de* $\mathfrak{A}$ *engendrée par l'ensemble* $A$.

La nécessité a déjà été prouvée. La suffisance découle de l'égalité $J_\mathfrak{A} = \mathscr{C}_{\mathscr{B}(\mathfrak{A})}$.

Décrivons maintenant les opérateurs de fermeture qui sont à la fois topologiques et algébriques.

THÉORÈME 7. *Soit* $P$ *un opérateur de fermeture sur un ensemble* $S$. *Les cinq conditions suivantes sont équivalentes*:

*) On trouvera la définition rigoureuse (par récurrence) du terme et de sa valeur en logique.

1) *P est un opérateur de fermeture topologique et algébrique;*
2) *P préserve les réunions arbitraires;*
3) *P est confondu avec l'opération de passage à l'image d'un ensemble par une relation $F_P$ sur S réflexive, transitive (unique):*

$$P(A) = F_P[A], \forall A \subset S;$$

4) *P est confondu avec l'opérateur générateur d'algèbres $J_{\mathfrak{A}}$ d'une algèbre $\mathfrak{A}$ sur S ne possédant que des opérations unaires dans sa signature.*
5) *Le système de parties fermées* Fix (*P*) *associé à l'opérateur de fermeture P est stable pour les réunions arbitraires.*

On effectue la démonstration de ce théorème par des raisonnements simples, en utilisant les faits déjà cités sur les opérateurs de fermeture, ainsi que l'équivalence des deux propositions suivantes, établie au n° 1.1 pour toute application $\varphi$ de l'ensemble $\mathfrak{P}(S)$ dans lui-même:
1) $\varphi$ préserve des réunions quelconques,
2) $\varphi$ est représentable sous forme d'une opération de passage à l'image par une coorrespondance $F: \varphi = F[\cdot]$.

Les espaces topologiques dans lesquels un système de parties fermées est stable pour des réunions arbitraires sont appelés *espaces de Kolmogorov* par N. Bourbaki dans « Topologie générale ». La relation $F_P$ mentionnée dans le n° 3 du théorème 3 est une relation de préordre sur $S$ définie en fonction de l'opérateur de fermeture de la manière suivante:

$$xF_Py \Leftrightarrow (x, y) \in F_P \Leftrightarrow y \in P\{x\}, \quad \forall x, y \in S. \rule{1pt}{1em}$$

Signalons que les opérateurs de fermeture sur $S$ qui se ramènent à une opération de passage à l'image d'ensembles pour une relation de préordre sur $S$, trouvent des applications fécondes en particulier depuis les travaux de S. Kripke, en logique mathématique, dans l'étude des logiques non classiques et notamment des systèmes modaux, temporels et intuitionnistes. Pour cette raison, les opérateurs de fermeture de cette nature seront appelés *opérateurs de Kripke.*

Soient $\mathfrak{A} = (A, \mathscr{A})$ et $\mathfrak{B} = (B, \mathscr{B})$ deux algèbres de même signature $\Omega$. Le produit direct $\mathfrak{A} \times \mathfrak{B}$ de ces algèbres (de même type) est défini de façon standard comme une algèbre de support $A \times B$ et d'opérations $\varphi^{(\mathfrak{A} \times \mathfrak{B})}$ qui se ramènent à une réalisation des opérations correspondantes $\varphi^{\mathfrak{A}}$ et $\varphi^{\mathfrak{B}}$ dans les algèbres facteurs. La correspondance $(A, B, F)$ dont le graphe $F$ est un ensemble stable pour toutes les opérations du produit $\mathfrak{A} \times \mathfrak{B}$ des algèbres sera appelée $\Omega$-*correspondance* entre les algèbres $\mathfrak{B}$ et $\mathfrak{B}$. On aurait pu visiblement définir la $\Omega$-correspondance entre $\mathfrak{A}$ et $\mathfrak{B}$ comme une sous-algèbre du produit $\mathfrak{A} \times \mathfrak{B}$ des algèbres $\mathfrak{A}$ et $\mathfrak{B}$.

THÉORÈME 8. *Soient* $\mathfrak{A} = (A, \mathcal{A})$ *et* $\mathfrak{B} = (B, \mathcal{B})$ *deux algèbres de signature* $\Omega$. *L'image* $F[M]$ *de tout ensemble* $M \subset A$ *stable pour les opérations de* $\mathfrak{A}$, *par la correspondance* $F$ *entre* $\mathfrak{A}$ *et* $\mathfrak{B}$ *dont le graphe est un ensemble* $F$ *stable pour les opérations de l'algèbre produit* $\mathfrak{A} \times \mathfrak{B}$, *sera aussi un ensemble stable pour les opérations de l'algèbre* $\mathfrak{B}$, *autrement dit l'image d'une sous-algèbre de* $\mathfrak{A}$ *par une* $\Omega$*-correspondance entre les algèbres* $\mathfrak{A}$ *et* $\mathfrak{B}$ *sera une sous-algèbre de* $\mathfrak{B}$.

Un exemple d'application de ce théorème est accessible au n° 1.6.

1.4. Comme signalé à la fin du n° 1.2, les parties d'un ensemble ordonné stables pour l'opération (partielle) inf ne sont pas nécessairement des systèmes d'éléments fermés (de points fixes) d'une fermeture sur cet ensemble ordonné. De ce point de vue la théorie des fermetures sur des ensembles ordonnés ne peut être interprétée comme une généralisation naturelle de la théorie des opérateurs de fermeture sur un ensemble, que l'on a étudiés au n° 1.3. Cependant on peut se tirer d'affaire en renonçant à la définition partout des fermetures et en considérant les fermetures partielles sur des ensembles ordonnés (c'est-à-dire les fermetures sur certaines parties de ces ensembles ordonnés). Si à tout ensemble $M$ on associe une fermeture partielle (que nous désignerons par $\mathcal{C}l\,(M)$ ou brièvement par $\mathcal{C}_M$) définie par:

$$x \mapsto \inf \{y \in M : x \leqslant y\} = \inf (M \cap [x, \rightarrow)),$$

alors seront stables pour l'opération inf les ensembles et eux seuls qui sont confondus avec les systèmes des éléments fermés des fermetures partielles de la forme $\mathcal{C}_M$, $M \subset X$, c'est-à-dire avec les ensembles de la forme Fix $(\mathcal{C}_M)$, où $M \subset X$.

Dans ce numéro on cite quelques propriétés générales des correspondances Fix et $\mathcal{C}l$ entre les fermetures partielles sur un ensemble ordonné et les parties de cet ensemble.

On dit qu'un ensemble $M$ d'un ensemble ordonné $\mathbf{X} = (X, \leqslant)$ est *stable* pour l'opération inf (ou encore pour l'intersection) et on note inf-stable si la borne inférieure (sous réserve qu'elle existe) de toute partie de $M$ est élément de $M$, autrement dit

$$(\forall A \subset M)\ (\forall\, y \in X)\ [y = \inf A \Rightarrow y \in M].$$

Supposons que $M \subset X$. On appellera *inf-enveloppe* de $M$ dans $\mathbf{X}$ le plus petit (au sens de l'inclusion) des ensembles inf-stables $A$ contenant $M$. Désignons par $[M]_{\inf}$ l'inf-enveloppe de $M$. Signalons que l'intersection de toute famille d'ensembles inf-stables est un ensemble inf-stable, et l'opération inf-enveloppe, c'est-à-dire $[\cdot]_{\inf}$, est évidemment un opérateur de fermeture sur $X$.

L'opération inf étant associative, l'inf-enveloppe $[M]_{\inf}$ de l'ensemble $M$ est manifestement composée des bornes inférieures de toutes les parties $A$ de $M$ pour lesquelles est définie inf $A$.

Dans la suite, on verra que les parties $A$ figurant dans la description de $[M]_{\inf}$ peuvent être d'une forme spéciale, ce qui précisera la structure de l'inf-enveloppe. Nous aurons besoin du lemme suivant sur le «graphe» de l'opération inf.

LEMME 1. *Soient* $A \subset X$, $y \in X$. *Si* $y = \inf A$, *alors*

$$y = \inf A_y = \inf (A \cap [y, \rightarrow)).$$

*Autrement dit*

$$\inf A \rightleftarrows \inf (A \cap [\inf A, \rightarrow)) \textit{ pour tous les } A \subset X.$$

Le lemme suivant est fondamental.

LEMME 2 (sur la structure de l'inf-enveloppe). *Soit* $M \subset X$. *Alors*

$$(\forall y \in X)\ [y \in [M]_{\inf} \Leftrightarrow y = \inf M_y = \inf (M \cap [y, \rightarrow))].$$

DÉMONSTRATION. L'implication $\Leftarrow$ est triviale. Prouvons l'implication $\Rightarrow$. Supposons que $y = \inf A$, où $A \subset M$. Montrons que $y = \inf (M \cap [y, \rightarrow))$. Il est clair que $y \in \beta^-(M \cap [y, \rightarrow))$, puisque $M \cap [y, \rightarrow) \subset [y, \rightarrow)$. Supposons maintenant que $z \in \beta^-(M \cap$ $\cap [y, \rightarrow))$ et prouvons que $z \leqslant y$. En effet, comme $A \subset M$, on a $A \cap [y, \rightarrow) \subset M \cap [y, \rightarrow)$ et par suite $z \in \beta^-(A \cap [y, \rightarrow))$. Or $y = \inf A$, donc d'après le lemme 1 $y = \inf (A \cap [y, \rightarrow))$. Par définition de inf il vient $y \leqslant z$, c.q.f.d.

Considérons maintenant pour tout $M \subset X$ la fonction $\mathscr{C}_M$ définie sur l'ensemble $M_*$ de tous les $x \in X$ pour lesquels est définie $\inf M_x = \inf (M \cap [x, \rightarrow))$:

$$\mathscr{C}_M(x) = \inf M_x = \inf (M \cap [x, \rightarrow)) = \inf \{y \in M : x \leqslant y\}, \ \forall x \in M_*$$

Du lemme 2 on déduit sans peine le théorème suivant sur la structure de l'inf-enveloppe en termes de fonctions $\mathscr{C}_M$.

THÉORÈME 9. *La fonction* $\mathscr{C}_M$, *où* $M$ *est un ensemble quelconque d'un ensemble ordonné* $\mathbf{X} = (X, \leqslant)$, *est une fermeture partielle sur* $\mathbf{X}$ *dont l'ensemble des éléments fermés est l'inf-enveloppe de* $M$:

$$[M]_{\inf} = \mathrm{Fix}(\mathscr{C}_M) = \mathscr{R}(\mathscr{C}_M) = \mathscr{C}_M[X].$$

DÉMONSTRATION. L'égalité précédente est une conséquence immédiate du lemme 1. Montrons que $\mathscr{C}_M$ est une fermeture partielle. L'égalité $\mathrm{Fix}(\mathscr{C}_M) = \mathscr{R}(\mathscr{C}_M)$ équivaut à l'idempotence de la fonction $\mathscr{C}_M$ définie sur $M_*$. D'où il résulte à propos que $[M]_{\inf} \subset$ $\subset M_* = \mathscr{D}(\mathscr{C}_M)$. Il est évident que $\mathscr{C}_M$ est croissante et dilatante sur son domaine de définition. Donc, $\mathscr{C}_M$ est une fermeture partielle sur l'ensemble ordonné $\mathbf{X}$.

On appellera $\mathscr{C}_M$ fermeture partielle engendrée par une base d'éléments fermés de $M$. L'opération $M \mapsto \mathscr{C}_M$, $M \subset X$, sera notée $\mathscr{C}l$. De sorte que $\mathscr{C}l(M) = \mathscr{C}_M$, $\forall M \subset X$.

Les fermetures partielles ne sont pas toutes de la forme $\mathscr{C}_M$. Le théorème et la remarque suivants situent la place occupée par les fermetures partielles de la forme $\mathscr{C}_M$ parmi les fermetures partielles sur X.

THÉORÈME 10. *Toute fermeture partielle $P$ sur un ensemble ordonné $(X, \leqslant)$, dont le système des éléments fermés est un ensemble $F$, est la restriction de la fermeture partielle $\mathscr{C}_F$:*

$$P = \mathscr{C}_F \mid \mathscr{D}(P).$$

*De plus, $\mathscr{C}_M = \mathscr{C}_N \Longleftrightarrow [M]_{\inf} = [N]_{\inf}$ pour tous les $M, N \subset X$ et en particulier $\mathscr{C}_M = \mathscr{C}_{[M]_{\inf}}$ pour tout $M \subset X$.*

DÉMONSTRATION. Prouvons la première partie du théorème. Soit $P$ une fermeture partielle sur $(X, \leqslant)$, telle que $\mathscr{R}(P) = \text{Fix}(P) = F$. Comme $P$ est une fermeture sur $(\mathscr{D}(P), \leqslant)$, elle est restaurée de façon unique par l'ensemble $\mathscr{R}(P) = F$: plus exactement

$$(\forall x \in \mathscr{D}(P)) \quad [P(x) = \text{Min}\, F_x].$$

Donc

$$(\forall x \in \mathscr{D}(P)) \quad [P(x) = \inf F_x].$$

Par définition de la fonction $\mathscr{C}_M$

$$(\forall x \in \mathscr{D}(P)) \quad [P(x) = \mathscr{C}_M(x)],$$

c'est-à-dire $P = \mathscr{C}_M/\mathscr{D}(P)$. Le reste du théorème découle de façon triviale de la première partie et du théorème 9. ▶

Les fermetures partielles $\mathscr{C}_M$, $M \subset X$, occupent une place très importante parmi les fermetures dont le domaine de définition $\mathscr{D}(\mathscr{C}_M)$ est fixe, et qui laissent invariants tous les points fixes des fermetures $\mathscr{C}_M$. Notamment, pour toute fermeture partielle $P$ telle que $\mathscr{D}(P) = \mathscr{D}(\mathscr{C}_M) = M_*$ et $\text{Fix}(\mathscr{C}_M) \subset \text{Fix}(P)$, on a

$$(\forall x \in M_*) \quad [\mathscr{C}_M(x) \geqslant P(x)], \quad \text{i.e.} \quad \mathscr{C}_M \geqslant P.$$

Ce qui vient d'être dit nous suggère la terminologie suivante. Les fermetures partielles sur $(X, \leqslant)$ de la forme $\mathscr{C}_M$, où $M \subset X$, seront dites *fermetures partielles satureés*. On appellera *saturé* d'une fermeture partielle $P$ sur $(X, \leqslant)$ la fermeture partielle saturée $\mathscr{C}_{\text{Fix}(P)} = \mathscr{C}_{\mathscr{R}(P)}$ et l'on notera $[P]_{\text{st}}$.

La fonction Fix qui associe aux fermetures partielles sur $(X, \leqslant)$ les ensembles de leurs éléments fermés, et la fonction $\mathscr{C}l$ qui associe aux parties de $X$ les fermetures partielles admettant ces parties pour base d'éléments fermés, établissent une correspondance entre $\mathfrak{P}(X)$ et toutes les fermetures partielles sur $(X, \leqslant)$, de telle sorte que

(1) $\text{Fix}(\mathscr{C}l(M)) = [M]_{\inf}$,

(2) $\mathscr{C}l(\text{Fix}(P)) = [P]_{\text{st}}$,

(3) $[\mathscr{C}l\ (M)]_{st} = \mathscr{C}l\ ([M]_{inf})$,
(4) $[\mathrm{Fix}\ (P)]_{inf} = \mathrm{Fix}\ ([P]_{st})$,
$M$ étant une partie quelconque de $X$, $P$ une fermeture partielle quelconque sur $(X, \leqslant)$. Les propriétés (1) et (2) résultent respectivement du théorème 9 et de la définition de $[\cdot]_{st}$. Les propriétés (3) et (4) sont une conséquence purement algébrique des propriétés (1) et (2).

On se propose maintenant de déduire un théorème sur les fonctions de la forme $\mathscr{C}_M$ pour le cas de treillis quasi complets.

Théorème 11. *Supposons que $(X, \leqslant)$ est un treillis quasi complet, qu'à tout ensemble $M \subset X$ est associé l'ensemble*

$$[M]_0 = \{x \in X;\ M_x = M \cap [x, \rightarrow) \neq \varnothing\}$$

*et que $\Pi_M$ est définie sur $[M]_0$ de la façon suivante:*

$$\Pi_M\ (x) = \inf M_x,\ \ \forall x \in [M]_0.$$

*Alors la fonction $\Pi_M$ est une fermeture sur l'ensemble ordonné $([M]_0, \leqslant)$ et prend ses valeurs sur l'ensemble $M_{inf}$ des bornes inférieures de toutes les parties non vides de $M$. L'ensemble $M_{inf}$ s'appelle enveloppe inférieure ou extension perronienne inférieure de $M$.*

Pour établir ce théorème comme corollaire du théorème 9 il suffit simplement de remarquer que la fonction $\Pi_M$ est la restriction d'une fonction $\mathscr{C}_M$ et de plus que ces deux fonctions sont confondues si l'ensemble ordonné $(X, \leqslant)$ ne possède pas d'élément maximum; si $(X, \leqslant)$ possède un élément maximum (qui sera visiblement $\inf \varnothing$), alors

$$\mathscr{C}_M\ (x) = \inf \varnothing,\ \ \forall x \in M_* - [M]_0.$$

Selon le théorème 7,1 opération $M \mapsto [M]_0$, c'est-à dire $[\cdot]_0$, est une fermeture sur $(\mathfrak{P}(X), \subset)$ qui préserve les réunions, c'est-à-dire une fermeture du type de Kripke, puisqu'elle est confondue avec le passage à l'image réciproque par la relation réflexive et transitive $\leqslant$ sur $X$:

$$[M]_0 = \geqslant [M],\ \ \forall M \subset X.$$

1.5. On se servira ici des polaires. Soit $F$ une correspondance d'un ensemble $X$ vers un ensemble $Y$.

On appelle *polaire* de l'ensemble $A \in X$ par la correspondance $F$ et on note $\pi_F\ (A)$ l'ensemble

$$\{y \in Y : (\forall x \in A)\ [(x,\ y) \in F]\}.$$

Donc, $y \in \pi_F\ (A)$ si et seulement si $(x,\ y) \in F$ pour tout $x \in A$. Il est évident que

$$\pi_F\ (A) = \{u \in Y : A \times \{u\} \subset F\}. \tag{1}$$

D'où il suit entre autres que si le polaire $\pi_F\ (A)$ n'est pas vide, alors $A \subset \mathscr{D}\ (F)$. Par ailleurs, si $A = \{x\}$, où $x \in X$, le polaire

$\pi_F(A)$ que nous désignerons brièvement par $\pi_F\{x\}$ est confondu avec l'image $F\{x\}$ de l'ensemble $\{x\}$ par $F$. Le polaire $\pi_{F^{-1}}(B)$ d'un ensemble $B \subset Y$ par $F^{-1}$ sera appelé aussi *polaire inverse* (de l'ensemble $B$ par la correspondance $F$) et désigné par $\pi_{F^{-1}}(B)$. Si aucune confusion n'est à craindre, on omettra d'écrire $F$ et $F^{-1}$ en indice.

De la définition il suit immédiatement que

I. *Si les ensembles $A$ est $B$ sont tels que $A \times B \subset F$, alors*

$$B \subset \pi(A),\ A \subset \pi^{-1}(B). \tag{2}$$

Comme de toute évidence $A \times \pi(A) \subseteq F$ et $\pi^{-1}(B) \times B \subseteq F$ pour tout ensemble $A \subset X$, la proposition I nous dit que

$$A \subset \pi^{-1}(\pi(A)),\ \forall A \subset X \tag{3}$$

et

$$B \subset \pi(\pi^{-1}(B))\ (B \subset Y). \tag{4}$$

De la définition du polaire il résulte en outre que

II. *La relation $A_1 \subset A_2 \subset X$ entraîne $\pi(A_1) \supset \pi(A_2)$. De même, la relation $B_1 \subset B_2 \subset Y$ entraîne $\pi^{-1}(B_2) \subset \pi^{-1}(B_1)$.*

Un sous-ensemble $A$ de $X$ n'est pas forcément polaire inverse d'un ensemble $B \subset Y$. Le critère suivant nous renseigne sur cette question.

THÉORÈME 12. *Soient $F$ une correspondance de $X$ vers $Y$, $A$ une partie de $X$. Pour qu'il existe un ensemble $B \subset Y$ tel que $A = \pi^{-1}(B)$ il est nécessaire et suffisant que pour tout élément $x \in X \setminus A$ on puisse exhiber un élément $y \in Y$ tel que*

$$A \subset \pi^{-1}\{y\},\ x \notin \pi^{-1}\{y\}. \tag{5}$$

*Si cette condition est remplie, alors $A = \pi^{-1}(\pi(A))$.*

DÉMONSTRATION. *La condition* est *nécessaire*. Si $B$ est une partie de $Y$ et $x$ un élément de $X$, tels que $A = \pi^{-1}(B)$ et $x \notin A$, alors par définition du polaire le produit $\{x\} \times B$ n'est pas contenu dans $F$, de sorte qu'il existe $y \in B$ pour lequel $(x, y) \notin F$. Donc $x \notin F^{-1}\{y\} = \pi^{-1}\{y\}$. Mais d'après la proposition II, $y \in B$ entraîne que $\pi^{-1}\{y\} \supset \pi^{-1}(B) = A$.

La *condition* est *suffisante*. Supposons que la condition du théorème est réalisée et prouvons que $A \supset \pi^{-1}(\pi(A))$, c'est-à-dire, compte tenu de (2), que $A = \pi^{-1}(\pi(A))$. Soit $x \in \pi^{-1}(\pi(A))$. Si, en outre, $x \notin A$, en choisissant un élément $y \in Y$ d'après le théorème, on aura en vertu de la proposition II:

$$y \in \pi(\pi^{-1}\{y\}) \subset \pi(A),$$

de sorte que $\pi^{-1}\{y\} \supset \pi^{-1}(\pi(A))$ et par suite $x \notin \pi^{-1}(\pi(A))$. ▶

Un ensemble $K \subset X$ s'appelle *F-composante* (ou simplement *composante* si $F$ est fixe) si $K = \pi^{-1}(\pi(K))$. Le théorème prouvé nous dit que le polaire inverse $\pi^{-1}(B)$ de tout ensemble $B \subset Y$ est une

$F$-composante. De façon analogue, le polaire $\pi(A)$ de tout ensemble $A \subset X$ est une $F^{-1}$-composante. De ce qui précède il résulte en particulier que le *bipolaire* $\pi^{-1}(\pi(A))$ d'un ensemble quelconque $A \subset X$ est une $F$-composante. D'après les propositions II et III l'application

$$\Pi_F : A \mapsto \pi^{-1}(\pi(A)) \quad (A \subset X) \tag{6}$$

est une fermeture sur l'ensemble ordonné $\mathfrak{W} = \mathfrak{P}(X)$ de toutes les parties de $X$, pour laquelle l'ensemble $\mathfrak{K}_F(X)$ de toutes les $F$-composantes est un système d'éléments fermés, qui de ce fait est un $\bigcap$-système de parties de $X$ (cf. 1.2).

Pour les mêmes raisons, l'application

$$\Pi_{F^{-1}} : B \mapsto \pi(\pi^{-1}(B)) \quad (B \subset Y) \tag{7}$$

sera une fermeture sur l'ensemble $\mathfrak{P}(Y)$ de toutes les parties de $Y$, dont le système d'éléments fermés est l'ensemble $\mathfrak{K}_{F^{-1}}(Y)$ de toutes les $F^{-1}$-composantes. Comme le polaire $\pi(K)$ de toute $F$-composante $K$ est une $F^{-1}$-composante et le polaire inverse $\pi^{-1}(L)$ d'une $F^{-1}$-composante $L$, une $F$-composante, les applications

$$\pi : K \mapsto \pi(K) \quad (K \in \mathfrak{K}_F(X))\,; \quad \pi^{-1} : L \mapsto \pi^{-1}(L) \quad (L \in \mathfrak{K}_{F^{-1}}(Y)) \tag{8}$$

sont inverses l'une de l'autre et, en vertu de la proposition II, sont des isomorphismes inverses des ensembles ordonnés $\mathfrak{K}_F(X)$ et $\mathfrak{K}_{F^{-1}}(Y)$.

La proposition suivante est une conséquence immédiate de la définition du polaire.

III. *Quel que soit l'ensemble $\mathfrak{A}$ des parties d'un ensemble $X$ on a la relation*

$$\pi(\bigcup \mathfrak{A}) = \bigcap_{A \in \mathfrak{A}} \pi(A). \tag{9}$$

Tout ensemble étant la réunion de ses parties à un seul élément, il s'ensuit que

$$\pi(A) = \bigcap_{x \in A} \pi\{x\} = \bigcap_{x \in A} F\{x\} \quad (A \subset X).$$

L'application $\Phi : x \mapsto \Pi_F\{x\}$ $(x \in X)$ qui plonge l'ensemble $X$ dans l'ensemble $\mathfrak{K}_F(X)$ de toutes les $F$-composantes s'appelle *injection canonique* (de l'ensemble $X$ dans l'ensemble $\mathfrak{K}_F(X)$). Les $F$-composantes couvrant le domaine de valeurs de l'application $\Phi$, c'est-à-dire les $F$-composantes de la forme $\Pi_F\{x\} = \pi^{-1}(\pi\{x\})$ $(x \in X)$, sont dites *principales*. L'ensemble $\mathfrak{K}_F(X)$ de toutes les $F$-composantes est généré en bas par l'ensemble de toutes les $F$-composantes principales : toute $F$-composante $K$ est borne supérieure (dans $\mathfrak{K}_F(X)$) de la famille $\{\Phi(x)\}$ $(x \in K) : K = \sup \Phi\,[K]$.

**1.6.** Considérons enfin un espace vectoriel X sur $\mathbb{K}$ *) et soit $\Lambda$ une correspondance de $\mathbb{K}$ vers $\mathbb{K}$, autrement dit $\Lambda$ est un sous-ensemble de $\mathbb{K}^2 = \mathbb{K} \times \mathbb{K}$. On dit qu'un ensemble $E \subset$ X est un $\Lambda$-*ensemble* si pour tous scalaires $\alpha$ et $\beta$ tels que $(\alpha, \beta) \in \Lambda$ on a $\alpha E + \beta E \subset \subset E$ **).

Donc, $E$ est un $\Lambda$ ensemble s'il contient deux quelconques de ses éléments $x$ et $y$ avec leur combinaison linéaire $\alpha x + \beta y$, $\alpha$ et $\beta$ étant des coefficients vérifiant la condition $(\alpha, \beta) \in \Lambda$. L'ensemble de tous les éléments de X sera un $\Lambda$-ensemble pour tout $\Lambda$.

I. *Quelle que soit la correspondance* $\Lambda$ *de* $\mathbb{K}$ *vers* $\mathbb{K}$ *l'ensemble* $\vartheta_\Lambda$ *de tous les* $\Lambda$-*ensembles est un* $\bigcap$-*système* ***) *de parties de l'espace* X.

En effet, si $\mathfrak{A}$ est un ensemble de $\Lambda$-ensembles, $E = \bigcap \mathfrak{A}$, $(\alpha, \beta) \in \in \Lambda$, pour tout $A \in \mathfrak{A}$ on a $\alpha E + \beta E \subset \alpha A + \beta A \subset A$, si bien que $\alpha E + \beta E \subset \bigcap \mathfrak{A} = E$.

La fermeture (ou le projecteur en haut) $\Pi_\Lambda$ de l'ensemble ordonné $\mathfrak{W} = \mathfrak{P}$ (X) sur le $\bigcap$-système $\vartheta_\Lambda$ de tous les $\Lambda$-ensembles associe à l'ensemble $E \subset$ X sa $\vartheta_\Lambda$-enveloppe $\Pi_\Lambda(E)$ que nous appellerons simplement $\Lambda$-*enveloppe* de $E$. D'après la définition générale, la $\Lambda$-enveloppe $\Pi_\Lambda(E)$ de l'ensemble $E$ est le plus petit de tous les $\Lambda$-ensembles contenant $E$.

D'une façon générale, l'ensemble $\vartheta_\Lambda$ de tous les $\Lambda$-ensembles n'est pas stable pour la réunion. Cependant

II. *Si* $\mathfrak{A}$ *est un ensemble filtrant supérieurement de* $\Lambda$-*ensembles, alors* $\bigcup \mathfrak{A}$ *est un* $\Lambda$-*ensemble.*

En effet, soient $x, y \in \bigcup \mathfrak{A}$, $(\alpha, \beta) \in \Lambda$. Cherchons des ensembbes $A_1, A_2 \in \mathfrak{A}$ tels que $x \in A_1$, $y \in A_2$. L'ensemble $\mathfrak{A}$ étant filtrant supérieurement par hypothèse, il existe dans $\mathfrak{A}$ un ensemble $A$ contenant aussi bien $A_1$ que $A_2$. Comme de toute évidence $x, y \in A$, on peut écrire: $\alpha x + \beta y \in \alpha A + \beta A \subset A \subset \bigcup \mathfrak{A}$.

Si l'espace vectoriel X est muni d'une topologie compatible avec sa structure, c'est-à-dire d'une topologie pour laquelle sont continues les opérations algébriques, alors il est aisé de comprendre que pour tout $\Lambda$ l'adhérence d'un $\Lambda$-ensemble est un $\Lambda$-ensemble. Nous nous trouvons donc dans les conditions du théorème 4 qui dit que toute application associant à un ensemble $e \subset$ X sa $\Lambda$-*adhérence*, c'est-à-dire l'adhérence $\mathscr{C}l\,(\Pi_\Lambda(e))$ de la $\Lambda$-enveloppe de $e$, est un projecteur

---

*) On rappelle que le corps des scalaires $\mathbb{K}$ est soit le corps $\mathbb{R}$ des réels, soit le corps $\mathbb{C}$ des complexes.

**) $A$, $B$, $C$ étant des ensembles dans l'espace vectoriel X et $\alpha$ un scalaire, par $A + B$ on désigne l'ensemble de tous les éléments de la forme $x + y$ $(x \in A,\ y \in B)$ et par $\alpha C$ l'ensemble de tous les éléments de la forme $\alpha z$ $(z \in C)$.

***) C'est-à-dire, un système $\bigcap$-stable.

en haut (i.e. une fermeture sur $(\mathfrak{P}(X), \subset)$). A noter que la $\Lambda$-enveloppe d'un ensemble fermé $e$, c'est-à-dire $(\mathscr{Cl} \circ \Pi_\Lambda)(e)$, n'est fermée que dans des cas exceptionnels. Donc $\mathscr{Cl} \circ \Pi_\Lambda$ et $\Pi_\Lambda \circ \mathscr{Cl}$ sont en général différentes l'une de l'autre.

Citons les classes de $\Lambda$-ensembles les plus fréquemment étudiées. Certaines de ces classes ont déjà été introduites et ne sont mentionnées que pour compléter le tableau.

1. $\Lambda = \mathbb{K}^2$. Dans ce cas un $\Lambda$-ensemble est un *ensemble linéaire*. Il est immédiat de montrer par récurrence que la combinaison linéaire $\sum_{k=1}^{n} \alpha_k x_k$ des éléments $x_k$ $(k = 1, 2, \ldots, n)$ d'un ensemble linéaire appartient à cet ensemble quels que soient les coefficients $\alpha_k$ $(k = 1, 2, \ldots, n)$. On en déduit sans peine que la $\Lambda$-enveloppe d'un ensemble $E$ (on l'appelle *enveloppe linéaire* et on la note $\mathscr{L}(E)$) est composée de tous les éléments de la forme $\sum_{k=1}^{n} \alpha_k x_k$, où $n$ est un entier naturel, $x_k \in E$ $(k = 1, 2, \ldots, n)$, $\alpha_k \in \mathbb{K}$ $(k = 1, 2, \ldots, n)$.

2. $\Lambda = \{(\alpha, \beta) \in \mathbb{K}^2 \colon \alpha + \beta = 1\}$. Les $\Lambda$-ensembles sont appelés *ensembles affines*. La $\Lambda$-enveloppe d'un ensemble $E \subset X$ s'appelle *enveloppe afine* de $E$ et se note Af $(E)$. Un ensemble affine est caractérisé par le fait qu'il contient avec $n$ quelconques de ses éléments $x_1, x_2, \ldots, x_n$ leur combinaison linéaire $\sum_{k=1}^{n} \alpha_k x_k$, les coefficients $\alpha_k$ vérifiant la relation $\sum_{k=1}^{n} \alpha_k = 1$. Il s'ensuit aussitôt que l'enveloppe affine Af $(E)$ d'un ensemble $E \subset X$ est composée de toutes les combinaisons linéaires $\sum_{k=1}^{n} \alpha_k x_k$ des éléments $x_1, x_2, \ldots, x_n \in E$, les coefficients $\alpha_k$ étant reliés par la relation $\sum_{k=1}^{n} \alpha_k = 1$, $n = 1, 2, \ldots$

3. $\Lambda = \{(\alpha, \beta) \in \mathbb{K}^2 \colon |\alpha| \leqslant 1, \beta = 0\}$. Les $\Lambda$-ensembles correspondant à un tel $\Lambda$ sont dits *ensembles équilibrés* et leurs enveloppes, *enveloppes équilibrées*. Il est clair que l'enveloppe équilibrée d'un ensemble $E \subset X$ est confondue avec la réunion $\bigcup_{|\alpha| \leqslant 1} \alpha E$.

4. $\Lambda = \{(\alpha, \beta) \in \mathbb{K}^2 \colon |\alpha| + |\beta| \leqslant 1\}$. Les $\Lambda$-ensembles correspondants sont dits *ensembles absolument convexes* et leurs $\Lambda$-enveloppes, *enveloppes absolument convexes*. Si $A$ est un ensemble absolument convexe, $x_1, x_2, \ldots, x_n \in A$ et les scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ tels que $\sum_{k=1}^{n} |\alpha_k| \leqslant 1$, alors $\sum_{k=1}^{n} \alpha_k x_k \in A$. On démontre que l'enveloppe absolument convexe d'un ensemble $E \subset X$ est composée de

tous les éléments de la forme $\sum_{k=1}^{n} \alpha_k x_k$, où $x_1, x_2, \ldots, x_n \in E$ et les scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ vérifient la condition: $\sum_{k=1}^{n} |\alpha_k| \leqslant 1$.

Les deux exemples suivants se rapportent au cas où **X** est un espace vectoriel réel (bien que les définitions formelles puissent être étendues à un espace complexe).

5. $\Lambda = \{(\alpha, \beta) \in \mathbf{R}^2 : \alpha, \beta \geqslant 0\}$. Un $\Lambda$-ensemble correspondant à un tel $\Lambda$ s'appelle *cône* (*convexe*) et sa $\Lambda$-enveloppe, *enveloppe conique*. Le cône contient toutes les combinaisons linéaires possibles de ses éléments avec des coefficients non négatifs quelconques. L'enveloppe conique d'un ensemble $E \subset \mathbf{X}$ est composée de toutes les combinaisons linéaires possibles des éléments de $E$ avec des coefficients non négatifs.

6. $\Lambda = \{(\alpha, \beta) \in \mathbf{R}^2 : \alpha, \beta \geqslant 0, \alpha + \beta = 1\}$. Un $\Lambda$-ensemble correspondant à un tel $\Lambda$ s'appelle *ensemble convexe* et sa $\Lambda$-enveloppe, *enveloppe $\Lambda$-convexe*. Comme dans les exemples précédents, on démontre qu'un ensemble convexe contient toutes les combinaisons linéaires de ses éléments, dont les coefficients sont non négatifs et la somme égale à un. L'enveloppe convexe d'un ensemble $E \subset \mathbf{X}$ est constituée de toutes les combinaisons possibles des éléments de $E$ avec des coefficients non négatifs de somme égale à un.

Nous ne nous arrêterons pas sur les innombrables relations liant les classes de $\Lambda$-ensembles associés à divers $\Lambda$. Signalons toutefois une relation évidente. Si $\Lambda_1 \subset \Lambda_2$, alors le $\cap$-système $\vartheta_{\Lambda_1}$ de tous les $\Lambda_1$-ensembles contient le $\cap$-système $\vartheta_{\Lambda_2}$ de tous les $\Lambda_2$-ensembles, si bien que par exemple un ensemble linéaire est un $\Lambda$-ensemble pour tout $\Lambda$, un ensemble absolument convexe est un ensemble équilibré, etc.

A noter encore que l'ensemble $\{0\}$ constitué du zéro de l'espace **X** est un $\Lambda$-ensemble quel que soit $\Lambda$. Si le couple $(0, 0) \in \Lambda$, l'élément nul de l'espace **X** appartient à tout $\Lambda$-ensemble non vide, c'est-à-dire l'ensemble $\{0\}$ est le plus petit de tous les $\Lambda$-ensembles non vides. En conséquence de quoi, si $(0, 0) \in \Lambda$, on n'entendra par $\Lambda$-ensemble qu'un $\Lambda$-ensemble *non vide*. Cette convention est valable pour les exemples 1, 3, 4 et 5.

Considérons encore un espace vectoriel **Y** sur le même corps de scalaires $\mathsf{K}$ que **X**. Munissons le produit $\mathbf{X} \times \mathbf{Y}$ d'une structure d'espace vectoriel de manière canonique. Une correspondance $U$ de **X** vers **Y** est *linéaire* si U est un ensemble linéaire dans $\mathbf{X} \times \mathbf{Y}$. De façon générale: $\Lambda$ étant une correspondance de $\mathsf{K}$ vers $\mathsf{K}$, on dit qu'une correspondance $U$ de **X** vers **Y** est une *$\Lambda$-correspondance* si $U$ est un $\Lambda$-ensemble dans l'espace vectoriel $\mathbf{X} \times \mathbf{Y}$.

III. *Soient $\Lambda$ une correspondance de $\mathsf{K}$ vers $\mathsf{K}$, $U$ une $\Lambda$-correspondance de* **X** *vers* **Y**. *L'image $U[E]$ d'un $\Lambda$-ensemble $E$ quelconque de* **X** *est un $\Lambda$-ensemble de* **Y**.

En effet si $y, v \in U[E]$, $(\alpha, \beta) \in \Lambda$, alors en choisissant les éléments $x, u \in E$ tels que $(x, y), (u, v) \in U$, on aura :

$$(\alpha x + \beta u, \alpha y + \beta v) = \alpha \cdot (x, y) + \beta \cdot (u, v) \in \alpha U + \beta U \subset U.$$

De $\alpha x + \beta u \in E$ il s'ensuit que $\alpha y + \beta v \in U[E]$.

La correspondance réciproque $U^{-1}$ étant aussi une $\Lambda$-correspondance, la proposition III s'étend aux images réciproques par la $\Lambda$-correspondance.

D'autre part, X et Y étant des $\Lambda$-ensembles pour tout $\Lambda$, le domaine de définition $\mathscr{D}(U) = U^{-1}[\mathrm{Y}]$ et le domaine de valeurs $\mathscr{R}(U) = U[\mathrm{X}]$ d'une $\Lambda$-correspondance quelconque $U$ (de X vers Y) sont des $\Lambda$-ensembles dans X et respectivement dans Y.

A signaler que dans la proposition III on entendra toujours par $U$ une correspondance linéaire.

1.7. Nous allons clore ce paragraphe par une bibliographie sur les sujets exposés.

La théorie des correspondances est traitée avec plus de détails dans la *Théorie des ensembles*, chapitre I, § 3, de Bourbaki, ainsi que dans *Universal algebra*, chapitre I, de P. M. Cohn, New York, Evanston and London, 1965. L'essentiel du n° 1.2 sur les fermetures sur un ensemble ordonné figure dans Bourbaki, *Théorie des ensembles*, chapitre III, § 1, exercices 12 à 15, § 2, exercices 6, 7. On trouvera aussi dans cet ouvrage une méthode générale de construction d'une fermeture qui est dans un certain sens la plus proche de la composée de deux fermetures, qui en général n'est pas une fermeture.

Les opérateurs de fermeture sur un ensemble ordonné introduits par E. Moore sont étudiés assez profondément dans l'ouvrage suscité de P. M. Cohn, chapitre II, 1.5, chapitre VII, 2.

La majeure partie du n° 1.6 relative aux $\Lambda$-ensembles et aux $\Lambda$-correspondances pour les espaces vectoriels est une conséquence immédiate des propositions générales sur les opérateurs de fermeture, présentées au n° 1.3.

Les notions et faits du n° 1.4, § 1, concernant les correspondances entre les sous-ensembles inf-stables d'ensembles ordonnés et les fermetures partielles saturées sur ces ensembles ordonnés appartiennent à G. Syrkine (à l'exception du théorème 11 sur les treillis quasi complets).

Il convient de signaler que nous n'avons pas abordé ici la question plus délicate et plus complexe de savoir quel doit être un ensemble ordonné (pris en tant que tel et non comme un sous-ensemble d'un ensemble donné *a priori*) pour être isomorphe (en tant que système algébrique d'un certain type) au système des éléments fermés d'un ensemble ordonné muni d'une fermeture, appartenant à une classe donnée *a priori* d'ensembles ordonnés « assez bons » munis des fermetures, car cela aurait impliqué des méthodes de logique. Ces questions sont abondamment traitées dans les ouvrages de logique

mathématique. Les systèmes d'éléments fermés associés aux classes d'algèbres booléennes munies d'une fermeture topologique sont amplement caractérisés dans les travaux consacrés aux problèmes des interprétations (au sens de K. Gödel, A. Tarski) des extensions de la logique intuitionniste des prédicats dans les extensions de la logique modale des prédicats S4. L'ouvrage de H. Rasiowa, R. Sikorski *The mathematics of metamathematics*, 3[rd] edition, Warszawa, 1970, est spécialement consacré à ces questions. G. Syrkine donne une caractéristique fonctionnelle complète des systèmes d'éléments fermés associés aux algèbres booléennes munies d'une fermeture (arbitraire et non topologique) dans son travail *On logics associated with closure operators*, Contributed papers to the 5[th] International Congress of Logic, Methodology and Philosophy of Science, Canada, 1975. Ce travail contient les calculs des séquences du type calculs de G. Gentzen, aussi bien pour les algèbres booléennes munies d'une fermeture que pour les systèmes d'éléments fermés qui leur sont associés. D'une façon générale, les calculs des séquences de cette nature constituent un appareil assez fin pour la recherche plus « efficace » de la démonstration d'une classe naturelle d'assertions sur des objets considérés. Ce travail contient également un théorème sur la représentation des algèbres booléennes munies d'une fermeture dans les algèbres booléennes des parties d'un ensemble muni d'un opérateur de fermeture de E. Moore, ce qui généralise le théorème classique de la représentation des algèbres booléennes munies d'une fermeture topologique dans les algèbres booléennes des parties d'un espace topologique (cf. par exemple l'ouvrage suscité de H. Rasiowa et R. Sikorski, chapitre XI, 10.2). A ces questions on peut en principe rapporter le théorème classique de G. Birkhoff et O. Frink qui fournit un critère pour qu'un treillis complet soit isomorphe à un treillis complet des sous-algèbres d'une algèbre universelle. Les liaisons de Galois entre les ensembles ordonnés (cf. G. Birkhoff, *Lattice theory* (revized edition) New York, 1948 (repr. 1960) sont une généralisation des polaires et des polaires inverses. L'essentiel du n° 1.5 aurait pu être exposé de toute évidence en termes plus généraux.

## § 2. Dualité algébrique entre espaces vectoriels

Dans ce paragraphe on étudie le lien unissant deux espaces vectoriels sur le produit desquels est défini un produit scalaire, fonctionnelle bilinéaire non dégénérée (ou séparable), linéaire par rapport au premier argument et conjuguée linéaire par rapport au second. L'existence du produit scalaire permet d'introduire toute une série de correspondances entre les espaces considérés. En termes de ces correspondances on peut formuler d'intéressantes propriétés pour les espaces vectoriels considérés.

**2.1.** Soient X et Y des espaces vectoriels sur le corps des scalaires $\mathbb{K}$. On dit qu'une fonctionnelle $S: (x, y) \mapsto S(x, y)$ $((x, y) \in \in X \times Y)$ est un *produit scalaire* si

1) $S$ est linéaire par rapport au premier argument et conjuguée linéaire par rapport au second, c'est-à-dire

$$S(\alpha x_1 + \beta x_2, \lambda y_1 + \mu y_2) = \alpha\bar{\lambda} S(x_1, y_1) + \\ + \alpha\bar{\mu} S(x_1, y_2) + \beta\bar{\lambda} S(x_2, y_1) + \beta\bar{\mu} S(x_2, y_2),$$

$$(\alpha, \beta, \lambda, \mu \in \mathbb{K};\ x_1, x_2 \in X;\ y_1, y_2 \in Y);$$

2) la fonctionnelle $S$ n'est pas *dégénérée* (ou bien, est *séparable*): pour tous éléments $x \in X$ et $y \in Y$ non nuls on peut exhiber des éléments $v \in Y$ et $u \in X$ tels que $S(x, v) \neq 0$ et $S(u, y) \neq 0$.

Les espaces X et Y sur le produit desquels est défini le produit scalaire $S$ sont dits *duaux*. On note ceci par le symbole $X \underset{S}{\leftrightarrow} Y$ ou simplement $X \leftrightarrow Y$ *).

Pour des raisons d'ordre historique on omet généralement d'écrire le symbole $S$ pour noter le produit scalaire $S(x, y)$ des éléments $x \in X$ et $y \in Y$. Cependant, les parenthèses (. . .) étant très souvent sollicitées en analyse fonctionnelle, on opte pour un autre mode de désignation. Pour notre part, nous écrirons $S(x, y) = \langle x, y \rangle$. A noter que le symbole $\langle x, y \rangle$ est susceptible de prêter à confusion dans la mesure où il peut être rapporté à un autre couple d'espaces. Pour éviter ceci nous indiquerons le couple d'espaces correspondants à chaque fois que le contexte l'exigera.

La dualité $X \leftrightarrow Y$ entre les espaces X et Y engendre une dualité $Y \leftrightarrow X$ entre les espaces Y et X, appelée *adjointe* de la dualité donnée. La dualité adjointe est définie par le produit scalaire *adjoint* du produit scalaire donné:

$$\langle y, x \rangle^* = \overline{\langle x, y \rangle} \quad (y \in Y, x \in X). \tag{1}$$

Si les espaces X et Y sont réels, la formule (1) devient:

$$\langle y, x \rangle^* = \langle x, y \rangle.$$

Considérons deux espaces X et Y en dualité et fixons $v \in Y$. Soit la foncionnelle $f_v: x \mapsto \langle x, v \rangle$ $(x \in X)$. Il est immédiat que $f_v$ est une fonctionnelle linéaire sur X et qu'en vertu de la non-dégénérescence du produit scalaire $f_v = 0$ si et seulement si $v$ est l'élément

*) On définit encore la dualité entre les espaces vectoriels en remplaçant 1) par la propriété suivante:

$$S(\alpha x_1 + \beta x_2, \lambda y_1 + \mu y_2) = \alpha\lambda S(x_1, y_1) + \alpha\mu S(x_1, y_2) + \\ + \beta\lambda S(x_2, y_1) + \beta\mu S(x_2, y_2)$$

$$(\alpha, \beta, \lambda, \mu \in \mathbb{K};\ x_1, x_2 \in X;\ y_1, y_2 \in Y)$$

qui exprime la bilinéarité de $S$.

nul de Y. Donc, l'application $\Theta: y \to f_y$ ($y \in$ Y) est linéaire (en vertu de la première condition de la définition du produit scalaire) et met en correspondance biunivoque l'espace Y et l'espace L (X) de toutes les fonctionnelles linéaires définies sur X. Inversement, si Y est un ensemble linéaire *total* *) dans l'espace L (X), les espaces X et Y sont mis en dualité par le produit scalaire:

$$\langle x, y\rangle = y(x) \quad (x \in \mathbf{X},\ y \in \mathbf{Y}).$$

Signalons que l'application $\Theta: y \mapsto f_y$ ($y \in$ Y) est alors identique.

Si l'on remplace la dualité donnée par son adjointe Y $\leftrightarrow$ X, on peut d'après la procédure précédente définir pour chaque élément $u \in$ X la fonctionnelle $g_u: y \mapsto \langle u, y\rangle$ ($y \in$ Y) et l'application $\Theta^*$: $x \mapsto g_x$ ($x \in$ X) de l'espace X dans l'espace L (Y) des fonctionnelles linéaires sur Y. De ce qui précède il suit que

I. *Si* X *et* Y *sont des espaces vectoriels en dualité et si l'un d'eux est de dimension finie, l'autre le sera également et de plus* dim X = = dim Y.

◀ En effet. supposons que dim X est finie. L'espace L (X) de toutes les fonctionnelles linéaires définies sur X est aussi de dimension finie et dim L (X) = dim X. Mais on a vu que les espaces Y et L (X) étaient isomorphes. Donc, l'espace Y est de dimension finie et dim Y $\leqslant$ dim L (X) = dim X. En remplaçant la dualité X $\leftrightarrow$ Y par la dualité adjointe Y $\leftrightarrow$ X on obtient en vertu de ce qui a été démontré que dim X $\leqslant$ dim Y. ▶

2.2. La dualité X $\leftrightarrow$ Y entre les espaces vectoriels X et Y engendre de nombreuses correspondances de X vers Y.

Soit $A$ un ensemble quelconque de scalaires. Posons

$$F_A = \{(x, y) \in \mathbf{X} \times \mathbf{Y}: \langle x, y\rangle \in A\}. \qquad (2)$$

Le polaire $\pi_{F_A}(E)$ de l'ensemble $E \subset$ X s'appelle *A-polaire* (de l'ensemble $E$). Point n'est besoin de formuler la définition du *A-polaire inverse* d'un ensemble $G \subset$ Y.

De toutes les formes de l'ensemble $A$ nous allons en étudier plus ou moins en détail deux seulement: $A = \{0\}$ et $A = \{\alpha \in \mathbb{K}: |\alpha| \leqslant \leqslant 1\}$. Dans le premier cas le $A$-polaire d'un ensemble $E \subset$ X s'appelle *orthocomplément* de l'ensemble $E$ et se note $E^\perp$. On définit de même l'orthocomplément d'un ensemble $G \subset$ Y, que n'on note aussi $G^\perp$, en remplaçant la dualité donnée par son adjointe. Dans le second cas (c'est-à-dire lorsque $A = \{\alpha \in \mathbb{K}: |\alpha| \leqslant 1\}$) le $A$-polaire d'un ensemble $E \subset$ X sera simplement appelé *polaire* (ou, cf. proposition I, *polaire absolument convexe*) et désigné par $\pi(E)$ **).

*) On rappelle qu'un ensemble de fonctionnelles linéaires définies sur un espace vectoriel donné est *total* si l'unique élément sur lequel s'annulent ces fonctionnelles est l'élément nul.

**) On désigne souvent par $E^\circ$ le polaire absolument convexe de $E$.

Signalons, même si nous n'aurons pas à nous en servir pour la suite, deux cas d'ensembles $A$ assez fréquents, se rapportant au cas réel (éventuellement formellement au cas complexe). Premier cas: $A = ]-\infty, 1]$. Le $A$-polaire d'un ensemble $E$ est dit *polaire de Minkowski*. Deuxième cas: $A = [0, +\infty[$. Le $A$-polaire d'un ensemble $E$ s'appelle *cône adjoint* de l'ensemble $E$ (voir proposition I).

Le corps des scalaires $\mathbb{K}$ sera traité comme un espace vectoriel de dimension un sur $\mathbb{K}$.

Outre l'ensemble $A$ des scalaires, nous considérons encore un ensemble $\Lambda \subset \mathbb{K}^2$. Appelons $\overline{\Lambda}$ l'ensemble de tous les couples $(\overline{\alpha}, \overline{\beta}) \in \mathbb{K}^2$ de vecteurs unidimensionnels tels que $(\alpha, \beta) \in \Lambda$.

I. *Si l'ensemble $A$ de vecteurs unidimensionnels est un $\overline{\Lambda}$-ensemble dans l'espace* $\mathbb{K}$ (cf. 1.4), *alors le $A$-polaire de tout ensemble $E \subset$* X *est un $\Delta$-ensemble* (*dans l'espace* Y).

◀ Posons $G = \pi_{F_A}(E)$ et considérons des éléments $y, v \in G$. Si des scalaires $\alpha$ et $\beta$ sont tels que $(\alpha, \beta) \in \Lambda$, donc $(\overline{\alpha}, \overline{\beta}) \in \overline{\Lambda}$, alors

$$\langle x, \alpha y + \beta v\rangle = \overline{\alpha} \langle x, y\rangle + \overline{\beta} \langle x, v\rangle \in A$$

puisque $\langle x, y\rangle$, $\langle x, y\rangle \in A$, $\forall x \in E$, par définition du $A$-polaire et $A$ est un $\overline{\Lambda}$-ensemble. Ce que est équivalent à $\alpha y + \beta v \in \pi_{F_A}(E) = G$, c'est-à-dire le $A$-polaire $G$ de l'ensemble $E$ est un $\Lambda$-ensemble. ▶

En particulier, l'orthocomplément $E^{\perp}$ de tout ensemble $E \subset$ X est un ensemble linéaire dans Y. Le polaire absolument convexe $\pi(E)$ est un ensemble absolument convexe.

II. *Si $E$ est un ensemble de l'espace* X, *son polaire* (*absolument convexe*) $\pi(E)$ *est confondu avec son orthocomplément $E^{\perp}$*: $\pi(E) = E^{\perp}$.

◀ En effet, pour tout $n = 1, 2, \ldots$ l'appartenance $x \in E$ entraîne $nx \in E$. Donc, pour $y \in \pi(E)$, $x \in E$, on doit avoir $|\langle nx, y\rangle| \leqslant 1$, c'est-à-dire $|\langle x, y\rangle| \leqslant \frac{1}{n}$ $(n = 1, 2, \ldots)$, ce que n'est possible que si $\langle x, y\rangle = 0$, c'est-à-dire lorsque $y \in E^{\perp}$. Par suite, $\pi(E) \subset E^{\perp}$. La relation inverse $E^{\perp} \subset \pi(E)$ a manifestement lieu pour tout ensemble $E \subset$ X. ▶

Signalons la proposition suivante.

III. *Les orthocompléments d'un ensemble $E \subset$* X *et de son enveloppe linéaire $\mathscr{L}(E)$ sont confondus*: $E^{\perp} = (\mathscr{L}(E))^{\perp}$.

◀ Comme $E \subset \mathscr{L}(E)$, la proposition II de 1.6 nous dit que $E^{\perp} \supset (\mathscr{L}(E))^{\perp}$. La relation inverse résulte de ce que les éléments de l'enveloppe linéaire $\mathscr{L}(E)$ sont les combinaisons linéaires des

éléments de l'ensemble $E$, de sorte que si $y \in E^{\perp}$ et

$$z = \sum_{k=1}^{n} \lambda_k x_k, \text{ où } \lambda_k \in \mathbb{K}, \quad x_k \in E$$
$$(k = 1, 2, \ldots, n),$$

alors

$$\langle z, y \rangle = \sum_{k=1}^{n} \lambda_k \langle x_k, y \rangle = 0,$$

c'est-à-dire $y \in (\mathscr{L}(E))^{\perp}$. ▶

**2.3.** Poursuivons l'étude de la dualité entre les espaces X et Y. Soient $X_0$ une partie de l'espace X, $Y_0 = X_0^{\perp}$ son orthocomplément. Désignons par $\overline{Y}$ l'espace quotient $Y/Y_0$. Soient $x \in X_0$, $\overline{y} \in \overline{Y}$ et $y'$, $y''$ des éléments de la classe d'équivalence $\overline{y}$. Comme $y'' - y' \in Y_0 = X_0^{\perp}$, on a

$$\langle x, y' - y'' \rangle = 0,$$

de sorte que $\langle x, y' \rangle = \langle x, y'' \rangle$. Ceci définit sur le produit $X_0 \times \overline{Y}$ une fonctionnelle associant au couple $(x, \overline{y}) \in X_0 \times \overline{Y}$ le scalaire $\langle x, y \rangle$, où $y$ est un élément quelconque de la classe d'équivalence $\overline{y}$. Cette fonctionnelle est un produit scalaire sur $X_0 \times \overline{Y}$. Sa bilinéarité est évidente. On vérifie sans peine qu'elle n'est pas dégénérée : si $x \in X_0$ et $x \neq 0$, alors, le produit scalaire initial étant non dégénéré, il existe un élément $v \in Y$ tel que $\langle x, v \rangle \neq 0$. Si l'on entend par $\overline{v}$ la classe d'équivalence contenant l'élément $v$, on obtient une valeur non nulle pour le (nouveau) produit scalaire $\langle x, \overline{v} \rangle$. D'autre part, si $\overline{y} \neq 0$, aucun des éléments de la classe d'équivalence $\overline{y}$ n'appartient à l'ensemble $Y_0 = X_0^{\perp}$. Donc, si l'on considère un élément quelconque $y_0 \in \overline{y}$, on peut choisir $u \in X_0$ tel que $\langle u, \overline{y} \rangle = \langle u, y_0 \rangle \neq 0$.

Désignons par $\psi$ l'isomorphisme canonique de l'espace Y sur l'espace quotient $\overline{Y} = Y/Y_0$. Entre le produit scalaire défini sur $X_0 \times \overline{Y}$ et le produit scalaire initial on a la relation

$$\langle x, \overline{y} \rangle = \langle x, \psi(y) \rangle = \langle x, y \rangle \quad (x \in X_0, \ y \in Y). \tag{2}$$

Le produit scalaire défini met l'espace vectoriel $X_0 \subset X$ et l'espace quotient $\overline{Y} = Y/Y_0 = Y/X_0^{\perp}$ en dualité appelée *dualité quotient* par la dualité initiale.

On peut « compléter » la dualité $X_0 \leftrightarrow \overline{Y}$ en remplaçant $X_0$ par la *composante orthogonale* $X_0^{\perp\perp}$. Comme $(X_0^{\perp\perp})^{\perp} = (X_0^{\perp})^{\perp\perp} = X_0^{\perp} = Y_0$ (cf. 1.7), on obtient une dualité appelée *dualité quotient complète par la dualité initiale*.

L'orthocomplément de tout ensemble étant un ensemble linéaire, il n'est point besoin d'admettre que $X_0$ est linéaire lorsqu'on envisage la dualité $X_0^{\perp\perp} \leftrightarrow Y/X_0^{\perp}$.

La proposition I de 2.1 nous conduit au fait suivant.

I. *Soient* X *et* Y *des espaces vectoriels en dualité. Tout ensemble linéaire* $X_0$ *de dimension finie est une composante orthogonale, de sorte que* $X_0^{\perp\perp} = X_0$.

◀ Considérons la dualité quotient $X_0 \leftrightarrow Y/X_0^{\perp}$ correspondant à la dualité $X \leftrightarrow Y$, et la dualité quotient complète $X_0^{\perp\perp} \leftrightarrow Y/X_0^{\perp}$. La proposition I de 2.1 appliquée à ces deux dualités nous dit que le bipolaire $X_0^{\perp\perp}$ est fini et a la même dimension que l'ensemble $X_0$. Comme, de plus, $X_0^{\perp\perp} \supset X_0$ (cf. 1.6), de ce qui précède il s'ensuit que $X_0^{\perp\perp} = X_0$. ▶

Voici un corollaire utile de la proposition I.

II. *Soient donnés deux espaces vectoriels* X *et* Y *en dualité et un ensemble fini* $E$ *d'éléments de* X *linéairement indépendants. Si un élément* $x_0 \in X$ *n'appartient pas à l'enveloppe linéaire* $X_0 = \mathcal{L}(E)$ *de l'ensemble* $E$, *on peut exhiber dans l'espace* Y *un élément* $y_0$ *tel que*

$$\langle x_0, y_0 \rangle = 1, \quad \langle x, y_0 \rangle = 0, \quad (x \in E). \tag{3}$$

◀ Désignons par $Y_0$ l'orthocomplément de l'ensemble $E$ ou ce qui revient au même en vertu de la proposition III de 2.2, de l'enveloppe linéaire $\mathcal{L}(E) = X_0$. La proposition I nous apprend que $X_0 = X_0^{\perp\perp} = Y_0^{\perp}$. Donc, $x_0 \notin Y_0^{\perp}$ et par suite on peut exhiber dans $Y_0$ un élément $y_0$ tel que $\langle x_0, y_0 \rangle \neq 0$. En multipliant l'élément $y_0$ par un scalaire convenable on peut sans violer la relation $y_0 \in Y_0$ (proposition I de 2.2) faire en sorte que $\langle x_0, y_0 \rangle = 1$. ▶

La proposition II est généralement énoncée sous forme d'un théorème de biorthogonalisation.

Théorème 1. *Soient* $\{x_\xi\}$ $(\xi \in \Xi)$ *une famille* ***finie*** *d'éléments linéairement indépendants d'un espace vectoriel* X, Y *un espace vectoriel dual de* X. *Il existe une famille* $\{y_\xi\}$ $(\xi \in \Xi)$ *d'éléments de l'espace* Y ***biorthogonale*** *à la famille* $\{x_\xi\}$ $(\xi \in \Xi)$, *c'est-à-dire telle que*

$$\langle x_\xi, y_\eta \rangle = \begin{cases} 0 & (\xi \neq \eta), \\ 1 & (\xi = \eta) \end{cases} \quad (\xi, \eta \in \Xi). \tag{4}$$

2.4. Supposons comme précédemment que les espaces vectoriels X et Y sont en dualité. Supposons par ailleurs que sur X est définie une topologie localement convexe $\tau$. On dira que cette topologie est *compatible* avec la dualité $X \leftrightarrow Y$ si pour tout $y \in Y$ la fonctionnelle $f_y$ (cf. 2.1) est continue pour la topologie $\tau$. Une topologie (localement convexe) sur l'espace Y est *compatible* avec la dualité $X \leftrightarrow Y$ si elle l'est avec la dualité adjointe $Y \leftrightarrow X$. Il est immédiat de s'assurer que ceci traduit la continuité de toute fonctionnelle $g_x$ $(x \in X)$ (cf. 2.1).

Dans la suite, si l'on envisage une dualité entre des espaces vectoriels X et Y dont l'un est muni d'une topologie localement convexe, on admettra implicitement la compatibilité de cette topologie et de cette dualité.

On sait que pour définir une topologie localement convexe sur un espace vectoriel X, il suffit d'exhiber un *filtre* formé de voisinages fermés absolument convexes de 0 de l'espace X, qui soit une base du filtre formé de tous les voisinages de 0. *)

Un ensemble absolument convexe $V$ d'un espace vectoriel X n'est pas toujours voisinage de l'élément nul de X pour une topologie localement convexe sur X. Pour qu'il existe sur X une topologie localement convexe pour laquelle un ensemble (absolument convexe) donné $V \subset X$ soit voisinage de 0, il est nécessaire et suffisant que $V$ soit *absorbant*: c'est-à-dire pour tout $x \in X$ on peut exhiber un nombre $\varepsilon$ strictement positif tel que $\alpha x \in V$ dès que le scalaire $\alpha$ vérifie la relation $|\alpha| \leqslant \varepsilon$. Si l'on admet *a priori*, ce que nous ferons, que l'ensemble $V$ est absolument convexe, la condition de la définition précédente devient: $X = \bigcup_{\alpha \geqslant 0} \alpha V$.

Il est immédiat que l'intersection d'un nombre *fini* d'ensembles absorbants est un ensemble absorbant, de sorte que c'est un filtre sur $\mathfrak{P}$ (X).

Soit X un espace localement convexe. On appelle *tonneau* (ou *quasi-voisinage* de 0) un ensemble $V \subset X$ absolument convexe, fermé et absorbant. On a signalé à maintes reprises que le filtre des voisinages de 0 possède une base constituée de tonnaux.

Donc, si l'on désigne par $\mathfrak{J} = \mathfrak{J}$ (X) l'ensemble des tonneaux d'un espace localement convexe donné X et que l'on observe que l'intersection d'un nombre *fini* de tonneaux est un tonneau, de sorte que l'ensemble $\mathfrak{J}$ ordonné par l'inclusion est un ensemble filtrant inférieurement, alors on peut dire que la topologie sur X est définie par un filtre sur $\mathfrak{J}$, le filtre de tous les voisinages tonnelés de 0. A noter que le filtre des voisinages tonnelés de 0 ne comprend pas

---

*) On rappelle qu'un *filtre* dans un inf demi-treillis $W$ est un ensemble non vide $F \subset W$ contenant les bornes inférieures de toutes ses parties finies non vides et possédant de plus la propriété dite de *stabilité en haut*: l'ensemble $F$ contient avec chacun de ses éléments $w$ l'intervalle $[w, \rightarrow)$. Un sous-ensemble $B$ d'un filtre $F$ est une *base* de ce filtre si $F = \bigcup_{w \in B} [w, \rightarrow)$, c'est-à-dire si tout élément de $F$ peut être approché inférieurement par un élément de $B$. Si T est un espace topologique et si l'on prend pour $W$ l'ensemble filtrant $\mathfrak{P}$ (T) de toutes les parties de T, l'ensemble de tous les voisinages d'un point quelconque fixe de l'espace T nous fournit un exemple assez typique de filtre dans $\mathfrak{P}$ (T).

On sait que dans un espace localement convexe l'élément nul possède une base de filtre de voisinages constituée d'ensembles fermés absolument convexes. L'ensemble de tous les ensembles fermés absolument convexes étant un $\bigcap$-système et *a fortiori* un ensemble filtrant, on peut parler des filtres formés de voisinages de cette nature.

en général tous les tonneaux de $\mathfrak{J}$ dans l'espace X. En effet, il existe dans X des tonneaux qui ne sont pas voisinages de 0. Si le filtre des voisinages tonnelés est confondu avec l'ensemble $\mathfrak{J}$ des tonneaux, c'est-à-dire si chaque tonneau est voisinage de 0, l'espace X est dit *tonnelé*.

Lorsqu'on considère un ensemble absolument convexe $U$ dans un espace vectoriel X, on ne dispose pas de critères affirmant l'existence dans X d'une topologie pour laquelle $U$ serait un tonneau. Les deux propositions suivantes sont de ce fait importantes.

I. *Soient* X *et* Y *des espaces vectoriels en dualité. Si sur* X *est définie une topologie compatible avec la dualité* $\mathrm{X} \leftrightarrow \mathrm{Y}$, *alors le polaire inverse* $\pi^{-1}(G)$ *de tout ensemble* $G \subset \mathrm{Y}$ *est fermé dans l'espace* X.

◀ Posons comme précédemment $A = \{\alpha \in \mathbb{K} : |\alpha| \leqslant 1\}$. L'ensemble $A$ est fermé dans l'espace des scalaires $\mathbb{K}$. Donc, l'image réciproque $f_y^{-1}[A]$ qui visiblement est confondue avec le polaire inverse $\pi^{-1}\{y\}$ le sera également, puisque la fonctionnelle $f_y$ ($y \in$ $\in$ Y) est continue. Reste à remarquer que $\pi^{-1}(G) = \bigcap\limits_{y \in G} \pi^{-1}\{y\}$ (cf. 1.6). ▶

Jointe à la proposition I, la proposition suivante nous fournit un indice assez commode pour établir si est tonnelé le polaire inverse $U = \pi^{-1}(G)$ d'un ensemble $G \subset \mathrm{Y}$.

II. *Pour que le polaire inverse* $U = \pi^{-1}(G)$ *d'un ensemble* $G \subset \mathrm{Y}$ *soit un ensemble absorbant, donc un tonneau, il est nécessaire et suffisant que pour tout* $x \in \mathrm{X}$ *l'on ait*

$$p(x) = \sup_{y \in G} |\langle x, y \rangle| < +\infty. \tag{5}$$

*Sous cette condition, la fonctionnelle p définie sur* X *par* (5) *est une semi-norme. La boule de rayon r correspondant à cette semi-norme est confondue avec l'ensemble* $rU$.

◀ Supposons que l'ensemble $U$ est absorbant. Pour tout $x \in \mathrm{X}$ il existe par définition un nombre strictement positif $\varepsilon$ tel que $\varepsilon x \in U$. On peut donc écrire pour tout $y \in G$:

$$\varepsilon |\langle x, y \rangle| = |\langle \varepsilon x, y \rangle| \leqslant 1,$$

de sorte que

$$\sup_{y \in G} |\langle x, y \rangle| \leqslant \frac{1}{\varepsilon} < +\infty.$$

En particulier, si $x \in U$, on peut prendre $\varepsilon = 1$. Donc,

$$U \subset \{x \in \mathrm{X} : p(x) \leqslant 1\}.$$

Supposons maintenant que pour un $x \in \mathrm{X}$ on ait

$$t = p(x) = \sup_{y \in G} |\langle x, y \rangle| < +\infty.$$

Si un scalaire $\alpha$ est assez petit pour que $|\alpha t| \leqslant 1$, alors pour tout $y \in G$ on aura $|\langle \alpha x, y\rangle| = |\alpha| \, |\langle x, y\rangle| \leqslant |\alpha t| \leqslant 1$ et par suite $\alpha x \in \pi^{-1}(G) = U$, c'est-à-dire l'ensemble $U$ est absorbant. Si $t \leqslant 1$, on peut prendre $\alpha = 1$. On obtient alors $x \in U$ et par suite la relation $\{x \in \mathrm{X} : p(x) \leqslant 1\} \subset U$.

La fonctionnelle $f_y : x \mapsto \langle x, y\rangle$ $(x \in \mathrm{X})$ étant linéaire pour tout $y \in G$, la fonctionnelle $x \mapsto |f_y(x)|$ $(x \in \mathrm{X})$ est une semi-norme sur **X**. Si la condition (5) est remplie, il est immédiat de vérifier que l'enveloppe supérieure $x \mapsto \sup_{y \in G} |f_y(x)| = p(x)$ $(x \in \mathrm{X})$ de la famille indiquée de semi-normes sera aussi semi-norme. ▶

2.5. La dualité entre les espaces vectoriels **X** et **Y** permet de définir sur chacun d'eux une multitude de topologies dont les extrêmes, dans un certain sens, sont la topologie faible (sur **X**) et la topologie (*)-faible (sur **Y**).

I. *Soient* **X** *et* **Y** *des espaces vectoriels en dualité. Parmi les topologies localement convexes sur* **X** *compatibles avec cette dualité il en existe une faible. Pour base du filtre formé des voisinages de* **0** *pour cette topologie on peut prendre l'ensemble des polaires inverses de tous les ensembles finis de l'espace* **Y**.

◀ La réunion d'une famille finie d'ensembles finis étant un ensemble fini, l'ensemble des polaires inverses de tous les ensembles finis de **Y** est un ensemble filtrant inférieurement, donc il existe sur **X** une topologie localement convexe $\sigma$ pour laquelle le filtre formé des voisinages de **0** admet pour base l'ensemble indiqué.

Soit $y \in \mathrm{Y}$. Si $A = \{\alpha \in \mathbb{K} : |\alpha| \leqslant 1\}$ et si l'on remarque que pour tout $\varepsilon > 0$ l'image réciproque d'un $\varepsilon$-voisinage fermé de 0 dans $\mathbb{K}$ (c'est-à-dire l'ensemble $\varepsilon A$ de $\mathbb{K}$) par l'application $f_y$ est:

$$f_y^{-1}[\varepsilon A] = f_{\frac{1}{\varepsilon} y}^{-1}[A] = \pi^{-1}\left\{\frac{1}{\varepsilon} y\right\},$$

c'est-à-dire un voisinage de **0** pour la topologie $\sigma$, on peut affirmer que la fonctionnelle $f_y$ est continue (pour la topologie $\sigma$), c'est-à-dire la topologie $\sigma$ est compatible avec la dualité $\mathrm{X} \leftrightarrow \mathrm{Y}$.

Si $G$ est un ensemble fini dans l'espace **Y**, on a

$$U = \pi^{-1}(G) = \bigcap_{y \in G} \pi^{-1}\{y\} = \bigcap_{y \in G} f_y^{-1}[A].$$

Soit $\tau$ une topologie localement convexe sur **X**, compatible avec la dualité entre **X** et **Y**. Alors les images réciproques $f_y^{-1}[A]$ $(y \in G)$ sont voisinages de **0** pour la topologie $\tau$, puisque les fonctionnelles $f_y$ $(y \in G)$ sont continues (cf. 2.1). Donc, l'intersection $U$ sera voisinage de **0** pour cette topologie. D'où il résulte que la topologie $\sigma$ est plus faible que la topologie $\tau$. ▶

La topologie $\sigma$ s'appelle *topologie faible* de l'espace X, définie par la dualité entre les espaces X et Y. On la note $\sigma$ (X, Y). La dualité adjointe $\mathrm{Y} \leftrightarrow \mathrm{X}$ engendre sur l'espace Y une topologie faible $\sigma$ (Y, X) appelée aussi *topologie* (*)*-faible* sur l'espace Y. Pour base du filtre formé des voisinages de 0 pour la topologie (*)-faible on peut prendre l'ensemble des polaires de tous les ensembles *finis* dans l'espace X.

Soit $\tau$ une topologie localement convexe sur un espace vectoriel X. Pour que la topologie $\tau$ soit compatible avec la dualité entre l'espace X et un espace vectoriel Y, il est nécessaire et suffisant que la topologie $\tau$ soit plus forte que la topologie faible $\sigma$ (X, Y). La nécessité de cette condition découle de la proposition I, la suffisance, du fait que pour tout $y \in \mathrm{Y}$ la fonctionnelle $f_y$ qui, en vertu de la proposition I de 2.5, est continue pour la topologie $\sigma$ (X, Y) le sera *a fortiori* pour la topologie plus forte $\tau$.

Une *famille filtrante à droite* $\{x_\xi\}$ $(\xi \in \Xi)$ d'éléments de X converge faiblement, c'est-à-dire pour la topologie faible $\sigma$(X, Y), vers l'élément nul de X si et seulement si pour chaque élément $y \in \mathrm{Y}$ converge vers 0 la suite $\{\langle x_\xi, y\rangle\}$ $(\xi \in \Xi)$.

En effet, si $\{x_\xi\}$ $(\xi \in \Xi)$ converge faiblement vers l'élément nul de l'espace X, et $\varepsilon$ est un nombre strictement positif, alors à partir d'un certain indice $\xi_0$ on doit avoir, pour tout $\xi \in \Xi$, $x_\xi \in \pi^{-1}\left\{\frac{y}{\varepsilon}\right\}$, de sorte que $|\langle x_\xi, y\rangle| \leqslant \varepsilon$ et par suite $\langle x_\xi, y\rangle \underset{\xi \in \Xi}{\longrightarrow} 0$.

Inversement, si $\langle x_\xi, y\rangle \underset{\xi \in \Xi}{\longrightarrow} 0$ pour tout $y \in \mathrm{Y}$, alors quels que soient l'ensemble fini $G \subset \mathrm{Y}$ et les indices $\xi \in \Xi$ assez éloignés on aura $|\langle x_\xi, y\rangle| \leqslant 1$ $(y \in G)$, de sorte que pour ces indices, $x_\xi \in \pi^{-1}(G)$, ce qui traduit la faible convergence de la famille $\{x_\xi\}$ $(\xi \in \Xi)$ vers l'élément nul de l'espace X.

Dans le même ordre d'idées on démontre que dire qu'un ensemble $E \subset \mathrm{X}$ est *faiblement borné* *) revient à dire que

$$\sup_{x \in E} |\langle x, y\rangle| < +\infty \quad (y \in \mathrm{Y}).$$

La proposition II de 2.4 s'énonce encore:

II. *Pour que le polaire* $\pi$ $(E)$ *d'un ensemble* $E \subset \mathrm{X}$ *soit absorbant dans l'espace* Y *il est nécessaire et suffisant que* $E$ *soit faiblement borné.*

On dira qu'une dualité $\mathrm{X} \leftrightarrow \mathrm{Y}$ est *exhaustive* pour une topologie localement convexe $\tau$ définie sur X, si $\tau$ est compatible avec la dualité donnée et si de plus l'ensemble X* des fonctionnelles linéaires continues sur X n'est composé que de fonctionnelles de la forme

*) C'est-à-dire borné pour la topologie faible $\sigma$ (X, Y). On rappelle que dans un espace localement convexe X un ensemble $E$ est *borné* s'il est *absorbé* par un voisinage quelconque $U$ de l'élément nul de X, ce qui traduit l'existence d'un nombre strictement positif $\varepsilon$ tel que $\alpha E \subset U$ dès que $|\alpha| \leqslant \varepsilon$.

$f_y$ ($y \in$ Y), autrement dit si l'application $\Theta: y \mapsto f_y$ ($y \in$ Y) est une application de Y sur X*.

La dualité canonique entre un espace localement convexe X et son dual X*, définie par le produit scalaire $\langle x, f\rangle = f(x)$ ($x \in$ X, $f \in$ X*), est trivialement exhaustive pour la topologie de l'espace X.

La proposition suivante nous fournit un exemple important de dualité exhaustive.

III. *Toute dualité* X $\leftrightarrow$ Y *entre des espaces vectoriels* X *et* Y *est exhaustive pour la topologie faible* $\sigma$ (X, Y).

◀ Soit $f$ une fonctionnelle linéaire faiblement continue sur X. Si $A = \{\alpha \in K: |\alpha| \leqslant 1\}$, l'image réciproque $f^{-1}[A]$ sera un faible voisinage de l'élément nul de X, de sorte qu'on peut exhiber un ensemble fini $G \subset$ Y tel que $f^{-1}[A] \subset \pi^{-1}(G)$. Prouvons que les éléments de l'ensemble $G$ peuvent être considérés comme linéairement indépendants. En effet, soient $y_0, y_1, \ldots, y_n$ des éléments de $G$ et $\alpha_1, \alpha_2, \ldots, \alpha_n$ des scalaires, tels que $y_0 = \sum_{k=1}^{n} \alpha_k y_k$. Comme

$$|\langle x, y_0\rangle| = |\langle x, \sum_{k=1}^{n} \alpha_k y_k\rangle| \leqslant C \sum_{k=1}^{n} |\langle x, y_k\rangle|,$$

où $C$ est un nombre supérieur à $\sum_{k=1}^{n} |\alpha_k|$, en posant $z_k = \frac{y_k}{C}$, on obtient

$$\pi^{-1}\{y_0\} \supset \bigcap_{k=1}^{n} \pi^{-1}\{z_k\}.$$

Donc, en remplaçant dans l'ensemble $G$ les éléments $y_0, y_1, \ldots, y_n$ par les éléments $z_1, z_2, \ldots, z_n$ et en omettant l'élément $y_0$, on obtient un ensemble $G_0$ tel que

$$\pi^{-1}(G_0) = \bigcap_{y \in G_0} \pi^{-1}\{y\} \subset \bigcap_{y \in G} \pi^{-1}\{y\} = \pi^{-1}(G) \subset f^{-1}[A].$$

Le nombre des éléments de $G_0$ est inférieur d'une unité au moins à celui des éléments de $G$.

En poursuivant cette procédure on obtient en fin de compte un ensemble $G$ linéairement indépendant.

Utilisons le théorème de biorthogonalisation (cf. 2.3) pour construire une famille $\{x_1, x_2, \ldots, x_n\}$ d'éléments de X biorthogonale à la famille $\{y_1, y_2, \ldots, y_n\}$:

$$\langle x_j, y_k\rangle = \begin{cases} 0 & (j \neq k), \\ 1 & (j = k) \end{cases} \quad (j, k = 1, 2, \ldots, n).$$

Posons

$$\lambda_k = \bar{f}(x_k) \quad (k=1,\ 2,\ \ldots,\ n), \quad y = \sum_{k=1}^{n} \bar{\lambda}_k y_k.$$

Soit $x \in \mathrm{X}$. Convenons que

$$x' = x - \sum_{k=1}^{n} \beta_k x_k, \quad \text{où} \quad \beta_k = \langle x,\ y_k \rangle.$$

Comme

$$\langle x',\ y_j \rangle = \langle x,\ y_j \rangle - \sum_{k=1}^{n} \beta_k \langle x_k,\ y_j \rangle = 0,$$

on obtient

$$x' \in G^{\perp} \subset f^{-1}[A].$$

L'ensemble $G^{\perp}$ étant linéaire (cf. 2.2), pour tout $\alpha \in \mathrm{K}$ on aura

$$\alpha x' \in G^{\perp} \subset f^{-1}[A], \quad \text{i.e.} \quad |f(\alpha x')| \leqslant 1,$$

ce qui n'est possible que pour $f(x') = 0$. Donc

$$f(x) = f\Big(\sum_{k=1}^{n} \beta_k x_k + x'\Big) = \sum_{k=1}^{n} \lambda_k \beta_k = \sum_{k=1}^{n} \bar{\lambda}_k \langle x,\ y_k \rangle =$$
$$= \langle x,\ \sum_{k=1}^{n} \bar{\lambda}_k y_k \rangle = \langle x,\ y \rangle = f_y^r(x)$$

Ce qui prouve que $f = f_y$. ▶

Si une dualité $\mathrm{X} \leftrightarrow \mathrm{Y}$ est exhaustive pour une topologie localement convexe donnée de X, on peut exhiber un critère qui nous permet de dire si un ensemble donné de X est une composante (absolument convexe).

THÉORÈME 2. *Supposons qu'une dualité* $\mathrm{X} \leftrightarrow \mathrm{Y}$ *est exhaustive pour une topologie localement convexe* $\tau$ *de* X. *Pour qu'un ensemble* $E$ *dans* X *soit une composante (absolument convexe), autrement dit que* $E = \pi^{-1}(\pi(E))$, *il est nécessaire et suffisant que* $E$ *soit fermé et absolument convexe.*

DÉMONSTRATION. La *condition* est *nécessaire* même sous l'hypothèse que la topologie $\tau$ n'est que compatible avec la dualité (cf. proposition I de 2.2 et proposition I de 2.4).

Prouvons la *condition suffisante*. Pour montrer que l'ensemble $E$ est une composante (absolument convexe) nous allons nous servir du théorème 2 du § 1. Prenons un élément $x_0 \in \mathrm{X}$ tel que $x_0 \notin E$. L'ensemble $E$ étant fermé, il existe un voisinage $V$ de $0$ tel que $E \cap (x_0 + V) = \varnothing$. Soit l'ensemble $U = E + V$. L'ensemble $U$ est ouvert, puisque $U = \bigcup_{x \in E} (x + V)$. De plus, il est absolument

convexe, donc c'est un voisinage (ouvert) de $0$. Comme $V = -V$, on a $x_0 \notin U$ (sinon il existerait un élément $v \in V$ tel que $x_0 = x + v$, où $x \in E$; comme $x = x_0 - v \in x_0 - V = x_0 + V$ on aurait contrairement à l'hypothèse $x \in E \cap (x_0 + V) \neq \varnothing$.

L'ensemble $U$ en tant que voisinage absolument convexe de $0$ définit une semi-norme $p$ de sorte que pour tout nombre réel strictement positif $r$ on a

$$\{x \in X : p(x) < r\} \subset rU \subset \{x \in X : p(x) \leqslant r\}. \qquad (6)$$

Comme, en particulier, $x_0 \notin U$, on a $p(x_0) > 1$. Le théorème de Banach (cf. p. 80, tome 1) nous dit que sur $X$ on peut définir une fonctionnelle linéaire $f$ telle que

$$|f(x_0)| > 1, \quad |f(x)| \leqslant p(x) \quad (x \in X). \qquad (7)$$

Si, comme toujours, $A = \{\alpha \in K : |\alpha| \leqslant 1\}$, on déduit de (6), que $f^{-1}[A] \supset U$ et la dernière relation exprime la continuité de la fonctionnelle $f$, car celle-ci est linéaire et $U$ est un voisinage de $0$. Or, la dualité $X \leftrightarrow Y$ est exhaustive pour la topologie de l'espace $X$. Donc, la continuité de la fonctionnelle $f$ entraîne l'existence d'un élément $y \in Y$ tel que

$$f(x) = f_y(x) = \langle x, y\rangle \quad (x \in X).$$

Les relations (7) nous donnent

$$E \subset E + V = U \subset f^{-1}]A] = \pi^{-1}\{y\},$$

et comme

$$|f(x_0)| = |\langle x_0, y\rangle| > 1$$

on a $x_0 \notin \pi^{-1}\{y\}$. Selon la remarque qui suit le théorème 2 du § 1 on obtient $E = \pi^{-1}\pi(E)$).

COROLLAIRE 1. *Si la dualité* $X \leftrightarrow Y$ *est exhaustive pour la topologie* $\tau$ *de l'espace* $X$, *alors une condition nécessaire et suffisante pour qu'un ensemble quelconque absolument convexe de* $X$ *soit fermé est qu'il soit faiblement fermé.*

◀ En effet, d'après la proposition III un ensemble faiblement fermé est fermé pour la topologie $\tau$ même si celle-ci n'est que compatible avec la dualité. Si cette dualité est exhaustive pour $\tau$, le théorème prouvé nous dit alors qu'un ensemble $E \subset X$ fermé (pour la topologie $\tau$) est une composante absolument convexe. Or, d'après la proposition III, la dualité $X \leftrightarrow Y$ est exhaustive pour la topologie faible $\sigma(X, Y)$ donc, toujours en vertu du théorème prouvé, $E$ est faiblement fermé. ▶

Le corollaire 1 nous permet d'énoncer le théorème 2 sous la forme:

COROLLAIRE 2. *Soit donnée une dualité* $X \leftrightarrow Y$. *Pour qu'un ensemble* $E \subset X$ *soit une composante absolument convexe, il est nécessaire et suffisant qu'il soit absolument convexe et faiblement fermé.*

COROLLAIRE 3. *Soit donnée une dualité* $X \leftrightarrow Y$ *exhaustive pour une topologie* $\tau$. *Pour qu'un ensemble* $E \subset X$ *soit composante orthogonale, c'est-à-dire* $E = E^{\perp\perp}$, *il est nécessaire et suffisant qu'il soit linéaire et faiblement fermé.*

◀ La *condition nécessaire* est évidente.

Pour prouver la *condition suffisante*, considérons un ensemble $E \subset X$ linéaire faiblement fermé. L'ensemble $E$ étant absolument convexe, le théorème 2 nous dit que $E = \pi^{-1}(\pi(E))$. Reste à remarquer que le polaire d'un ensemble linéaire est confondu avec l'orthocomplément de cet ensemble. ▶

Soit $E$ un ensemble dans un espace vectoriel X. Par abs co $(E)$ on désigne son enveloppe absolument convexe (cf. 1.4) et par $\overline{E}^{\sigma}$ son adhérence faible, c'est-à-dire son adhérence pour la topologie faible $\sigma(X, Y)$ engendrée sur X par la dualité entre X et Y.

IV. *La composante absolument convexe engendrée par l'ensemble* $E \subset X$, *c'est-à-dire l'ensemble* $\pi^{-1}(\pi(E))$, *est confondue avec l'adhérence faible* $\overline{\text{abs co }(E)}^{\sigma}$ *de l'enveloppe absolument convexe* abs co $(E)$ *de* $E$.

◀ L'ensemble abs co $(E)$ est par définition absolument convexe. Donc, son adhérence faible $\overline{\text{abs co }(E)}^{\sigma}$ sera aussi absolument convexe (et faiblement fermée) (cf. 1.4). Par conséquent, $\overline{\text{abs co }(E)}^{\sigma}$ sera une composante absolument convexe contenant manifestement l'ensemble $E$ et par suite $\overline{\text{abs co }(E)}^{\sigma} \supset \pi^{-1}(\pi(E))$, puisque $E \subset \pi^{-1}(\pi(E))$. Inversement, $\overline{\text{abs co }(E)}^{\sigma} \subset \pi^{-1}(\pi(E))$ en raison de la convexité absolue et de la faible fermeture de la composante $\pi^{-1}(\pi(E))$ (corollaire 2 du théorème 2). ▶

Si au lieu du corollaire 2 du théorème 2 on se sert du corollaire 3 du même théorème on aboutit au résultat suivant.

V. *La composante orthogonale engendrée par l'ensemble* $E \subset X$, *c'est-à-dire l'ensemble* $E^{\perp\perp}$, *est confondue avec l'adhérence faible* $\overline{\mathscr{L}(E)}^{\sigma}$ *de l'enveloppe linéaire* $\mathscr{L}(E)$ *de l'ensemble* $E$.

On remarquera que si dans les conditions de la proposition IV l'ensemble $E$ est absolument convexe, et dans les conditions de la proposition V, linéaire, alors les composantes correspondantes seront simplement les adhérences de $E$. On notera encore que si l'espace X est muni d'une topologie pour laquelle est exhaustive la dualité $X \leftrightarrow Y$, alors dans les propositions IV et V et bien sûr dans les remarques qui les suivent on peut parler, en se rappelant la proposi-

tion III, des adhérences pour la topologie mentionnée (et non pour la faible).

La propriété d'être exhaustive pour une topologie est préservée par le passage de la dualité donnée $X \leftrightarrow Y$ à la dualité quotient (cf. 2.3).

VI. *Si une dualité* $X \leftrightarrow Y$ *est exhaustive pour une topologie* $\tau$ *sur* X, *alors pour toute partie* $X_0$ *de* X *la dualité quotient* $X_0 \leftrightarrow \overline{Y} = Y/Y_0$ $(Y_0 = X_0^{\perp})$ *est exhaustive pour une topologie* $\tau_0$ *de* $X_0$.

◀ Soit $f_0$ une fonctionnelle linéaire sur $X_0$, continue pour la topologie $\tau_0$. Soit par ailleurs $U$ un voisinage absolument convexe de l'élément nul de $X_0$, tel que $U \supset f_0^{-1}[A]$, où comme toujours $A = \{\alpha \in \mathbb{K} : |\alpha| \leqslant 1\}$. Choisissons un voisinage absolument convexe $V$ de l'élément nul de l'espace X, tel que $U = V \cap X_0$. L'ensemble $V$ en tant qu'ensemble absolument convexe et absorbant engendre sur X une semi-norme $p$ reliée au voisinage $V$ par la relation:

$$\{x \in X: p(x) < 1\} \subset V \subset \{x \in X: p(x) \leqslant 1\},$$
$$(\forall x \in X)\ [p(x) = \inf\{r > 0 : x \in rV\}]$$

Il s'ensuit que

$$(\forall x \in X_0)\ [|f(x)| < p(x)]$$

puisque $V$ est absorbant. Le théorème de Banach de prolongement des fonctionnelles linéaires affirme l'existence sur X d'une fonctionnelle linéaire $f$ qui prolonge $f_0$ et vérifie la relation $|f(x)| \leqslant p(x)$ $(x \in X)$. De cette relation il résulte que $f^{-1}[A] \supset V$, de sorte que la fonctionnelle $f$ est continue (pour la topologie $\tau$). La dualité entre X et Y étant exhaustive pour la topologie $\tau$, il existe dans l'espace Y un élément $y$ tel que $f(x) = f_y(x) = \langle x, y\rangle$ $(x \in X)$. En particulier pour $x \in X_0$ on aura

$$f(x) = f_y(x) = \langle x, y\rangle = \langle x, \overline{y}\rangle$$

où $\overline{y}$ est la classe d'équivalence de $y$. ▶

VII. *Si une dualité* $X \leftrightarrow Y$ *est exhaustive pour une topologie* $\tau$ *de l'espace* X *et si* $X_0$ *est une partie faiblement fermée de* X, *la dualité adjointe* $\overline{X} = X/X_0 \leftrightarrow X_0^{\perp}$ *de la dualité quotient* $X_0^{\perp} \leftrightarrow X/X_0^{\perp\perp} = \overline{X}$ *est exhaustive pour la topologie quotient* $\overline{\tau}$ *de l'espace quotient* $\overline{X} = X/X_0$.

◀ Désignons par $\varphi$ l'homomorphisme canonique de l'espace X sur l'espace quotient $\overline{X} = X/X_0$. Si par $Y_0$ on comprend l'orthocomplément $X_0^{\perp}$, alors d'après ce qui a été dit au 2.3 le produit scalaire définissant la dualité $\overline{X} \leftrightarrow Y_0$ est donné par

$$\langle \overline{x}, y\rangle = \langle \varphi(x), y\rangle = \langle x, y\rangle \quad (\overline{x} \in \overline{X},\ \overline{x} = \varphi(x),\ y \in Y_0). \qquad (8)$$

Soit $\bar{f}$ une fonctionnelle linéaire continue (pour la topologie $\bar{\tau}$) sur l'espace quotient $\bar{X}$. Si l'espace $\bar{X}$ est muni de la topologie $\bar{\tau}$ on sait alors que l'application $\varphi$ sera continue (pour la topologie $\tau$ de X) *). Posons $f = \bar{f} \circ \varphi$. Il est clair que la fonctionnelle $f$ est continue pour la topologie $\tau$ et par suite, la dualité $X \leftrightarrow Y$ étant exhaustive pour la topologie $\tau$, il existe un élément $y \in Y$ vérifiant la relation

$$f(x) = f_y(x) = \langle x, y \rangle \quad (x \in X). \tag{9}$$

La fonctionnelle $f$ s'annulant sur $X_0$, la relation (9) nous conduit à l'appartenance $y \in X_0^{\perp} = Y_0$, de sorte que, en vertu de (8), pour tout élément $\bar{x} \in \bar{X}$ on obtient

$$\bar{f}(\bar{x}) = f(x) = \langle x, y \rangle = \langle \varphi(x), y \rangle = \langle \bar{x}, y \rangle \quad (\varphi(x) = \bar{x}),$$

c.q.f.d. ▶

2.6. Soient X et Y des espaces vectoriels en dualité. On considérera l'application $\Theta$ qui à un élément $y \in Y$ associe la fonctionnelle $f_y : x \mapsto \langle x, y \rangle$ $(x \in X)$ comme une application de l'espace Y dans le produit $\mathbb{K}^X$ composé de toutes les fonctions scalaires (numériques) définies sur X. L'application $\Theta$ est linéaire, biunivoque et bicontinue si l'on munit l'espace Y de la topologie (*)-faible $\sigma(Y, X)$ et l'espace $\mathbb{K}^X$ de la topologie du produit. On établit sans peine la dernière propriété par exemple à l'aide de la remarque sur le caractère de la convergence faible. Donc, $\Theta$ est un homéomorphisme linéaire de l'espace Y sur l'espace $\mathbb{K}^X$.

I. *Si une dualité* $X \leftrightarrow Y$ *est exhaustive pour une topologie* $\tau$ *de l'espace* X, *alors le polaire* $\pi(V)$ *de tout voisinage* $V$ *de l'élément nul de* X *(pour la topologie* $\tau$*) est* (*)-*faiblement compact, autrement dit est compact dans l'espace* Y *muni de la topologie* $\sigma(Y, X)$.

◀ D'après ce qui a été dit sur l'application $\Theta$, il suffit d'établir la compacité dans l'espace $\mathbb{K}^X$ de l'image $\Theta[\pi(V)]$ que nous désignerons par $M$. En remplaçant, si cela est nécessaire, le voisinage $V$ par la composante absolument convexe qu'il engendre (le polaire $\pi(V)$ reste invariant par ce changement qui nous donne de nouveau un voisinage puisque $\pi^{-1}(\pi(V)) \supset V$, on peut au départ admettre que $V$ est une composante absolument convexe, autrement dit est un ensemble faiblement fermé et absolument convexe. Etant un voisinage de $\mathbf{0}$, $V$ est un ensemble absorbant et d'après ce qui a été dit précédemment: $V = \pi^{-1}(\pi(V))$. La proposition VI de 2.5 nous dit que le polaire $\pi(V)$ est faiblement borné, de sorte qu'en

*) La topologie quotient est précisément définie comme la plus forte des topologies de $\bar{X}$ pour lesquelles est continu l'homomorphisme canonique $\varphi$.

admettant que pour tous les $x \in X$

$$p(x) = \sup_{y \in \pi(V)} |\langle x, y\rangle|,$$

$$B_x = \{u \in \mathbb{K}: |u| \leqslant p(x)\},$$

on peut conclure que $M \subset \prod_{x \in X} B_x$. Or, pour tout $x$ l'ensemble $B_x$ qui est borné et fermé dans $\mathbb{K}$ est compact dans $\mathbb{K}$, donc, d'après le théorème de Tikhonov, le produit $\prod_{x \in X} B_x$ le sera également dans $\mathbb{K}^X$. Prouvons que l'ensemble $M$ est fermé dans l'espace $\mathbb{K}^X$. Soit $\varphi$ un point d'adhérence de l'ensemble $M$. Il existe une suite généralisée $F$ d'éléments de $M$ convergeant vers $\varphi$ dans l'espace $\mathbb{K}^X$. Autrement dit, pour chaque $x \in X$ la valeur $\varphi(x)$ est la limite d'une famille numérique $\{f(x)\}$ $(f \in F)$. Si l'on convient que $\Xi = \Theta^{-1}[F]$, en munissant l'ensemble $\Xi$ de l'ordre naturel (induit à partir de $F$), on peut écrire que $\varphi(x) = \lim_{\xi \in \Xi} \langle x, \xi\rangle$ $(x \in X)$. Il s'ensuit que la fonction $\varphi$ est linéaire. Comme $\Xi = \Theta^{-1}[F] \subset \pi(V)$, on a $|\varphi(x)| \leqslant 1$ pour $x \in V$, ce qui entraîne la continuité (pour la topologie $\tau$) de la fonction $\varphi$, puisque celle-ci est linéaire. Ainsi, $\varphi$ est une fonctionnelle linéaire continue (pour la topologie $\tau$) sur X. Mais la dualité $X \leftrightarrow Y$ est exhaustive pour la topologie $\tau$, donc $\varphi \in \Theta[Y]$. Vu que $|\varphi(x)| \leqslant 1$ $(x \in V)$, il vient $\Theta^{-1}(\varphi) \in \pi(V)$ et par suite $\varphi \in \Theta[\pi(V)] = M$. Étant un sous-ensemble fermé de l'ensemble compact $\prod_{x \in X} B_x$, l'ensemble $M$ est compact avec son image réciproque $\pi(V)$ par l'application $\Theta$.

La topologie de X pour laquelle la base de filtre des voisinages de 0 est constituée des polaires inverses de tous les ensembles absolument convexes (∗)-faiblement compacts de Y s'appelle *topologie de Mackey* sur X, engendrée par la dualité $X \leftrightarrow Y$. Cette topologie est notée $\mu(X, Y)$ *).

La dualité $X \leftrightarrow Y$ étant exhaustive pour la topologie faible $\sigma(X, Y)$, la proposition I nous dit que le polaire $\pi(V)$ de tout voisinage faible $V$ de 0 dans X est (∗)-faiblement compact. Si, en particulier, $V = \pi^{-1}(K)$, où $K$ est un ensemble fini dans Y, on obtient que la composante $C = \pi(\pi^{-1}(K)) = \pi(V)$ engendrée par $K$ est (∗)-faiblement compacte. Comme de plus $\pi^{-1}(K) = \pi^{-1}(C)$, l'ensemble $V = \pi^{-1}(K)$ est voisinage de 0 pour la topologie de Mackey $\mu(X, Y)$. Donc, cette topologie est plus forte que la topologie faible $\sigma(X, Y)$. Ce qui veut dire en particulier que la topologie de

*) Il est aisé de comprendre que si $C_1$ et $C_2$ sont des compacts dans un espace localement convexe, leur somme $C_1 + C_2$ le sera également. Si $C_1$ et $C_2$ sont des ensembles absolument convexes (∗)-faiblement compacts, leur somme le sera aussi. Il est évident que

$$\pi^{-1}(C_1) \cap \pi^{-1}(C_2) \supset \pi^{-1}(C_1 + C_2).$$

Mackey $\mu(X, Y)$ est compatible avec la dualité $X \leftrightarrow Y$, c'est-à-dire chaque fonctionnelle de la forme $f_y$ ($y \in Y$) est continue pour cette topologie.

Complétons la proposition I en prouvant que

II. *La dualité* $X \leftrightarrow Y$ *est exhaustive pour la topologie de Mackey* $\mu(X, Y)$.

◀ Outre Y, considérons l'espace $X_0'$ des fonctionnelles faiblement continues sur X. Il est clair que tous les éléments de cet espace sont des fonctionnelles de la forme $f_y$ ($y \in Y$). Désignons par ailleurs par $X_\mu'$ l'espace de toutes les fonctionnelles linéaires sur X, continues pour la topologie de Mackey $\mu(X, Y) = \mu$. La topologie $\mu$ étant plus forte que la topologie faible $\sigma(X, Y)$, on a de toute évidence $X_0' \subset X_\mu'$.

Considérons les dualités canoniques $X \leftrightarrow Y_\sigma'$ et $X \leftrightarrow X_\mu'$ *). Désignons le polaire et le polaire inverse respectivement par $\pi$ et $\pi^{-1}$ dans la dualité $X \leftrightarrow X_\sigma'$ et par $(\ldots)^\circ$ dans la dualité $X \leftrightarrow X_\mu'$. Il s'ensuit immédiatement de la définition que

$$\pi(E) = E^0 \cap X_\sigma' \quad (E \subset X) \tag{10}$$

et que $\pi^{-1}(G) = G^\circ$, pour $G \subset X_\sigma'$.

Prenons une fonctionnelle quelconque $f \in X_\mu'$ et supposons comme toujours que $A = \{\alpha \in K : |\alpha| \leqslant 1\}$. La fonctionnelle $f$ étant continue pour la topologie $\mu$, l'image réciproque $f^{-1}[A]$ est voisinage de l'élément nul de X pour cette topologie, de sorte qu'il existe un ensemble absolument convexe (*)-faiblement compact $C \subset X_\sigma'$ tel que

$$\{f\}^0 = f^{-1}[A] \supset \pi^{-1}(C) = C^0 \tag{11}$$

Mais la relation (10) nous dit que l'espace $X_\sigma'$ muni de la topologie faible $\sigma(X_\sigma', X)$ est un sous-espace de l'espace $X_\mu'$ muni de la topologie $\sigma(X_\mu', X)$. L'ensemble $C$ qui est compact dans le sous-espace mentionné est compact et *a fortiori* fermé dans l'espace $X_\mu'$ tout entier. Comme, d'autre part, $C$ est absolument convexe, on déduit que $C^{\circ\circ} = C$. En passant aux polaires dans (11) on peut donc écrire que

$$f \in \{f\}^{\circ\circ} \subset C^{\circ\circ} \subset C \subset X_\sigma'.$$

Par conséquent, $X_\mu' = X_\sigma'$, autrement dit toute fonctionnelle continue pour la topologie $\mu$ sur X est de la forme $f_y$ pour un $y \in Y$.

En combinant les propositions I et II on est conduit au fait central de la théorie de la dualité: le théorème de *Mackey-Arens*.

THÉORÈME 3. *Pour qu'une dualité* $X \leftrightarrow Y$ *soit exhaustive pour une topologie localement convexe* $\tau$ *donnée sur* X, *il est nécessaire et suffisant*

*) La première de ces dualités est visiblement isomorphe dans un sens à préciser à la dualité donnée, de sorte que $\sigma(X, X_\sigma') = \sigma(X, Y)$, $\mu(X, X_\sigma') = \mu(X, Y)$.

*que* $\tau$ *soit plus forte que la topologie faible* $\sigma$ (X, Y) *et plus faible que la topologie de Mackey* $\mu$ (X, Y).

*Nécessité.* Elle résulte de la proposition I.

*Suffisance.* Désignons par $X'_\sigma$ l'espace de toutes les fonctionnelles faiblement continues sur X, par X′ l'espace de toutes les fonctionnelles linéaires sur X, continues pour la topologie $\tau$ et enfin par $X'_\mu$ l'espace de toutes les fonctionnelles linéaires sur X, continues pour la topologie $\mu$ (X, Y). La relation entre les topologies $\sigma$ (X, Y), $\tau$ et $\mu$ (X, Y) nous permet d'écrire: $X'_\sigma \subset X' \subset X'_\mu$. Mais $X'_\mu = X'_\sigma$ d'après la proposition II. Donc $X' = X'_\sigma$. Reste à remarquer que l'égalité établie exprime justement que la dualité $X \leftrightarrow Y$ est exhaustive pour la topologie $\tau$. ▶

# BIBLIOGRAPHIE

## Ouvrages d'analyse fonctionnelle


AHIEZER N., GLAZMAN I.
Théorie des opérateurs linéaires dans l'espace hilbertien. Moscou, 1966 (en russe).

AKILOV G., MAKAROV B., HAVINE V.
Introduction élémentaire à la théorie de l'intégrale. Izd. LGU, 1969 (en russe).

ALEXANDROV P.
Introduction à la théorie générale des ensembles et des fonctions. M., Gostekhizdat, 1948 (en russe).

ANTONOVSKI M., BOLTIANSKI V., SARYMSAKOV T.
Algèbres topologiques de Boole. Izd. AN OuzSSR, 1963 (en russe).

BAKELMAN I.
Méthodes géométriques de résolution des équations elliptiques. M., « Naouka », 1965 (en russe).

BANACH S.
Théorie des opérations linéaires. Warszawa, 1932.

BESSOV O., ILYNE V., NIKOLSKI S.
Représentations intégrales des fonctions et théorèmes d'immersion. M., « Naouka », 1975 (en russe).

BIRKHOFF G.
Lattice Theory, N.Y., 1948.

BIRMAN M. et autres
Analyse fonctionnelle (dans la série « Spravotchnaya matematitcheskaya biblioteka »). M., « Naouka », 1972 (en russe).

BOURBAKI N.

I. Eléments de mathématique. Première partie. Livre I. Théorie des ensembles. Chapitre 1. Description de la mathématique formelle. Chapitre 2. Théorie des ensembles. Deuxième édition. Paris, Hermann, 1960.

II. Eléments de mathématique. Première partie. Les structures fondamentales de l'analyse. Livre I. Théorie des ensembles. Chapitre III. Ensembles ordonnés. Cardinaux. Nombres entiers. Paris, Hermann et C°, 1956.

III. Eléments de mathématique I. Les structures fondamentales de l'analyse. Livre I. Théorie des ensembles. Chapitre IV. Structures. Paris, Hermann, 1957.

IV. Eléments de mathématique I. Les structures fondamentales de l'analyse. Livre I. Théorie des ensembles. Fascicule des résultats. Troisième édition. Paris, Hermann, 1958.

V. Eléments de mathématique. Livre III. Topologie générale. Chapitre I. Structures topologiques. Chapitre II. Structures uniformes. Troisième édition. Paris, Hermann.

VI. Eléments de mathématique. Livre V. Espaces vectoriels topologiques. Paris, 1955.
VII. Eléments de mathématique. Première partie. Les structures fondamentales de l'analyse. Livre VI. Intégration. Mesures, intégration des mesures. Paris. Hermann et C°.
VIII. Eléments de mathématique. Première partie. Les structures fondamentales de l'analyse. Livre VI. Intégration. Intégration vectorielle, mesure de Haar, convolution et représentations. Paris.
IX. Théories spectrales. Chapitre I. Algèbres normés. Chapitre II. Groupes localement compacts. Paris, Hermann, 1967.
CHILOV G.
I. Analyse mathématique. Cours spécial. M., Fizmatguiz, 1961.
II. Analyse mathématique. Deuxième cours spécial. M., « Naouka », 1965.
CHILOV G., HUREWICZ B.
Intégrale, mesure et dérivée. M., « Naouka », 1964 (en russe).
COLLATZ L.
Funktionalanalysis und numerische Mathematik. B.R.D., 1964.
DAY M.
Normed Linear Spaces, Berlin, 1958.
DÉMIANOV V., MALOZÉMOV V.
Introduction au minimax. M., « Naouka », 1972 (en russe).
DÉMIANOV V., ROUBINOV A.
Méthodes approchées de résolution des problèmes d'extrémum. L., Izd-vo LGU, 1968 (en russe).
DIESTEL J.
Geometry of Banach Spaces., Selected topics, Lecture Notes in Math. 485, Springer-Verlag, Berlin-Heidelberg-New York, 1975.
DIEUDONNÉ J.
Foundations of Modern Analysis, N.Y., 1960.
DINCULEANU N.
Vector measures, Berlin, 1966.
DUNFORD N., SCHWARTZ J.
I. Linear Operators. General Theory, N.Y., 1958.
II. Linear Operators. Spectral theory, N.Y., 1963.
III. Linear Operators. Spectral Operators, N.Y., 1971.
EDWARDS R.
Functional Analysis. Theory and Applications, N.Y., 1965.
FADDÉEV D., FADDÉEVA V.
Méthodes numériques d'algèbre linéaire. M., Fizmatguiz, 1963 (en russe).
FIKHTENGOLTZ G.
Cours de calcul différentiel et intégral. Tomes I, II, III, M., « Naouka », 1966 (en russe).
FREMLIN D. H.
Topological Riesz spaces and measure theory, Cambridge Univ. Press, 1974.
GAVOURINE M.
Cours sur les méthodes de calcul. M., « Naouka », 1971 (en russe).
GLAZMAN I., LUBITCH Y.
Analyse linéaire dans les espaces à dimensions finies, Editions Mir, Moscou, 1972.
GOLDSTEIN E.
Théorie de la dualité en programmation mathématique et ses applications. M., « Naouka », 1971.
GUELFAND I., CHILOV G.
Les distributions. Editions Dunod.
Tome 1. Les distributions.
Tome 2. Espaces fondamentaux.
Tome 3. Théorie des équations différentielles.

GUELFAND I., RAÏKOV D., CHILOV G.
Anneaux normés commutatifs. M., Fizmatguiz, 1960 (en russe).
GUELFAND I., VILENKINE N.
Les distributions. Editions Dunod.
Tome 4. Applications de l'analyse harmonique.
HALMOS P.
Measure Theory, N.Y., 1950.
HARDY G., LITTLEWOOD J. a.o.
Inequalities, Cambridge, 1934.
HAUSDORFF F.
Théorie des ensembles. M., 1937.
HAVINE V.
Espaces de fonctions analytiques. V. zb. « Mathem. analiz. 1964 (Itogui naouki, VINITI AN SSSR) », M., 1966 (en russe).
HÖRMANDER L.
Linear Partial Differential Operators, N.Y., 1963.
HILBERT D.
Grundzüge einer allgemeine Theorie der Integralgleichungen, Leipzig, 1912.
HILLE E., PHILLIPS R.
Functional Analysis and Semi-Groups, Providens, 1957.
HOFFMAN K.
Banach Spaces of Analytic Functions, Englewood Cliffs, 1962.
IOFFE A., TIKHOMIROV V.
Théorie des problèmes d'extrémum. M., « Naouka », 1974.
KADETS M.
Géométrie des espaces normés. V. zb. « Mathem. analiz. 1975 (Itogui naouki. VINITI AN SSSR) ». M., 1975 (en russe).
KANTOROVITCH L., KRYLOV V.
Méthodes approchées d'analyse supérieure. M.-L., Fizmatguiz, 1962 (en russe).
KANTOROVITCH L., VOULIKH B., PINSKER A.
Analyse fonctionnelle dans les espaces semi-ordonnés. M.-L., Gostekhizdat, 1950 (en russe).
KARLIN S.
Mathematical Methods and Theory in Games, Programming and Economics, Vols. 1, 2, London, 1959.
KELLEY J.
Topologie générale. M., 1968.
KOLMOGOROV A., FOMINE S.
Eléments de la théorie des fonctions et de l'analyse fonctionnelle, Editions Mir, Moscou, 1974.
KOROTKOV V.
Opérateurs intégraux. Novossibirsk, izd-vo NGU, 1977 (en russe).
KOUTATELADZE S., ROUBINOV A.
La dualité de Minkowski et ses applications. Novossibirsk, « Naouka », 1976 (en russe).
KRASNOSSELSKI M.
I. Méthodes topologiques en théorie des équations intégrales non linéaires. M., Gostekhizdat, 1956 (en russe).
II. Solutions positives des équations opératorielles. M., Fizmatguiz, 1962 (en russe).
KRASNOSSELSKI M. et autres
Résolution approchée des équations opératorielles. M., « Naouka », 1969 (en russe).
KRASNOSSELSKI M. et autres
Opérateurs intégraux dans les espaces de fonctions sommables. M., « Naouka », 1966 (en russe).

KRASNOSSELSKI M., ROUTITSKI Y.
Fonctions convexes et espaces d'Orlicz. M., Fizmatguiz, 1958 (en russe.)
KRYLOV N.
Méthodes de résolution approchée des problèmes de physique mathématique. Izbrani troudy, t. 2, 150-204, Izd. AN USSR, Kiev, 1961 (en russe).
LADYJENSKAÏA O.
Problèmes aux limites de physique mathématique. M., « Naouka », 1973 (en russe).
LINDENSTRAUSS J., TZAFRIRI L.
Classical Banach spaces, Lecture Notes in Math., 338, Springer-Verlag, Berlin-Heidelberg-New York, 1973.
LUSTERNIK L., SOBOLEV V.
Eléments d'analyse fonctionnelle. M., « Naouka », 1965 (en russe).
LUXEMBURG W. A. J., ZAANEN A. C.
Riesz Spaces, v. I, Amsterdam, 1971.
MIKHLINE S.
I. Réalisation numérique des méthodes variationnelles. M., « Naouka », 1966 (en russe).
II. Méthodes variationnelles en physique mathématique. M., « Naouka », 1968 (en russe).
NAÏMARK M.
Anneaux normés. M., « Naouka », 1968 (en russe).
NAKANO H.
Modulared semi-ordered linear spaces, Tokyo, Maruzen Co., LTD, 1950.
NATANSON I.
I. Théorie constructive des fonctions. M., Gostekhizdat, 1949 (en russe).
II. Théorie des fonctions d'une variable réelle. M., « Naouka », 1974 (en russe).
NIKAÏDO H.
Convex Structures and Economic Theory, N.Y., 1968.
NIKOLSKI N.
Problèmes choisis d'approximation pondérée et d'analyse spectrale. Troudy Math. in-ta V. Stéklova, 120, « Naouka », L., 1974 (en russe).
PCHÉNYTCHNI B.
Conditions nécessaires d'extrémum. M., « Naouka », 1969 (en russe).
PÉTROVSKI I.
Cours sur la théorie des équations intégrales. M., « Naouka », 1965 (en russe).
PHELPS R.
Lectures on Choquet's Theorem. N.Y., 1966.
PHILIPPOV A., RIABENKI V.
Sur la stabilité des équations aux différences. M., Gostekhizdat, 1956 (en russe).
PLESNER A.
Théorie spectrale des opérateurs linéaires. M., « Naouka », 1965 (en russe).
RIESZ F., SZ.-NADY B.
Leçons d'analyse fonctionnelle, Budapest, 1953.
ROBERTSON A., ROBERTSON W.
Topological Vector Spaces, Cambridge, 1964.
ROCKAFELLAR R.
Convex Analysis, Princeton, 1970.
RUBINSTEIN G.
Modèles d'optimisation en dimension finie. Novossibirsk, « Naouka », 1970 (en russe).
RUDIN W.
Functional Analysis, N.Y., 1973.
SCHAEFER H.
I. Topological Vector Spaces, N.Y., 1966.

II. Banach lattices and positive operators, Springer-Verlag, Berlin-Heidelberg-New York, 1974.

SCHWARTZ L.
Théorie des distributions, t. I-II, Paris, 1950-1951.

SINGER I.
Bases, in Banach spaces, v. I, Springer-Verlag, Berlin-New York, 1970.

SMIRNOV V.
Cours de mathématiques supérieures, t. 5, M., Fizmatguiz, 1959 (en russe).

SOBOLEV S.
I. Quelques applications de l'analyse fonctionnelle à la physique mathématique. Izd-vo LGU, 1950 (en russe).
II. Introduction à la théorie des formules de cubature. M., « Naouka », 1974 (en russe).

SZEGÖ G.
Polynômes orthogonaux. M., Fizmatguiz, 1962 (en russe).

TIKHOMIROV V.
Quelques problèmes de la théorie des approximations. M., Izd-vo MGU, 1976 (en russe).

VAÏNIKKO G.
Approximation compacte des opérateurs et résolution approchée des équations. Tartou, Izd. Tartouskovo gos. un-ta, 1970 (en russe).

VLADIMIROV D.
Algèbre de Boole. M., « Naouka », 1969 (en russe).

VOROBIEV Y.
Méthode des moments en mathématiques appliquées. M., Fizmatguiz, 1958 (en russe).

VOULIKH B.
I. Introduction à la théorie des espaces semi-ordonnés. M., Fizmatguiz, 1961 (en russe).
II. Introduction à l'analyse fonctionnelle. M., « Naouka », 1967 (en russe).
III. Précis de la théorie des fonctions d'une variable réelle. M., « Naouka », 1973 (en russe).
IV. Introduction à la théorie des cônes dans les espaces normés. Izd. KGU, Kalinine, 1977 (en russe).

YOSIDA K.
Functional Analysis, Berlin, 1965.

ZAANEN A.
I. Linear Analysis, Amsterdam, 1960.
II. Integration, Amsterdam, 1967.

ZYGMUND A.
Trigonometrie Series, Vol. 1, Cambridge, 1959; Vol. 2, Cambridge, 1960.

## Ouvrages utilisés

ABRAMOVITCH Y.
1. Quelques théorèmes sur les espaces de Riesz normés. Vestnik LGU, nº 13 (1971), 5-11 (en russe).

AGAKHANOV S., NATANSON G.
1. Fonction de Lebesgue des sommes de Fourier-Jacobi. Vestnik LGU, nº 1 (1968), 11-23 (en russe).

AKILOV G.
1. Sur le prolongement des opérations linéaires. DAN SSSR 57, nº 7 (1947), 643-646 (en russe).
2. Conditions nécessaires de prolongement des opérations linéaires. DAN SSSR 59, nº 3 (1948), 417-418 (en russe).
3. Sur l'application d'une méthode de résolution des équations différentielles non linéaires à l'étude de systèmes d'équations différentielles. DAN SSSR 68, nº 4 (1949), 645 (en russe).

ALSYNBAEV K., IMOMNAZAROV B., RUBINSTEIN G.
I. Problème de déplacement d'une masse sur un compact muni d'une métrique non symétrique. V. sb. Optimizatsia, vyp. 17 (34), Novossibirsk, 1975, 94-107 (en russe).

ANDO T.
1. Linear functionals on Orlicz spaces. Nieuw Arch. Wiskunde 8, nº 1 (1960), 4-16.

ARONSZAIN N., SMITH K.
1. Invariant subspaces of completely continuous operators. Ann. Math. 60 (1954), 345-350.

BALOUEV A.
1. Sur la méthode de S. Tchapliguine. Vestnik LGU, nº 13 (1956), 27-42 (en russe).

BANACH S.
1. Sur les opérations dans les ensembles abstraits et leurs applications aux équations intégrales. Fund. Math. 3 (1922), 133-181.
2. Sur les fonctionnelles linéaires II. Studia Math. 1 (1929), 223-240.

BANACH S., STEINHAUS H.
1. Sur le principe de la condensation de singularités. Fund. Math. 9 (1927), 50-61.

BERKOLAÏKO M., RUTITSKI Y.
1. Sur les opérateurs dans les espaces généralisés de Helder, Sib. mathem. journal 12, 5 (1971), 1015-1025 (en russe).

BERMAN D.
1. Sur une classe d'opérations linéaires. DAN SSSR 85, 1 (1952), 13-16 (en russe).

BIRMAN M.
1. Quelques estimations pour la méthode de plus forte pente. UMN 5, nº 3 (1950), 152-155 (en russe).

2. Sur le calcul des valeurs propres par la méthode de plus forte pente. Zap. Léningrad. gornovo in-ta, 27, nº 1 (1952), 209-216 (en russe).

BIRMAN M., SOLOMIAK M.

1. Estimations des valeurs singulières des opérateurs intégraux. UMN 32, 1 (1977) 17-84 (en russe).

BOUKHVALOV A.

1. Sur les espaces localement convexes engendrés par des ensembles faiblement compacts. Vestnik LGU, nº 7 (1973), 11-17 (en russe).
2. Sur les espaces de norme mixte. Vestnik LGU, nº 19 (1973), 5-12 (en russe).
3. Sur la représentation intégrale des opérateurs linéaires. Zap. naoutch. séminarov Lening. otd. Mathem. in-ta AN SSSR, 47 (1974), 5-14 (en russe).
4. Opérateurs intégraux et représentation de fonctionnelles complètement linéaires sur des espaces de norme mixte. Sib. math. journ. 16, nº 3 (1975), 483-493 (en russe).
5. Sur la représentation analytique des opérateurs de norme abstraite. Izvestia vouzov, Mathem., nº 11 (1975), 21-32 (en russe).

BOUKHVALOV A., LOZANOVSKI G.

1. Sur les ensembles fermés pour la mesure dans des espaces de fonctions mesurables. Troudy Moskovskovo mathem. ob-va, 34 (1976), 128-149 (en russe).

BROUWER L. E. F.

1. Über Abbindung von Mannigfaltigkeiten, Math. Ann. 71 (1911), 97-115.

CURRY H. B.

1. The method of steepest descent for noulinear minimization problem. Quart. Appl. Math. 2, 3 (1944), 268-271.

DAOUGAVET I.

1. Application de la théorie générale des méthodes d'approximation à l'étude de la convergence de la méthode de Galerkine pour certains problèmes aux limites de physique mathématique. DAN SSSR 98, 6 (1954) (en russe).
2. Sur la méthode des moments pour une équation différentielle ordinaire. Sib. mathem. journ. 6, 1 (1965), 70-85 (en russe).

DOUBOVITSKI A., MILIOUTINE A.

1. Problèmes d'extrémum moyennant des contraintes. JVM et MF 5, nº 3 (1965), 395-453 (en russe).
2. Conditions nécessaires d'extrémum faible dans les problèmes de commande optimale avec des contraintes mixtes de type inégalités. JVM et MF 8, nº 4 (1968), 725-779 (en russe).

DUNFORD N., PETTIS B. J.

1. Linear operators on summable functions. Trans. Amer. Math. Soc. 47, nº 3 (1940), 323-392.

EKHLAKOV N.

1. Existence de transports optimaux de la mesure de Lebesgue avec densités. V zb. Optimizatsia, vyp. (15) (32), Novossibirsk, 1974, 90-114 (en russe).

ENFLO P.

1. A countreexample to the approximation problem in Banach spaces. Acta Math. 130, nº 3-4 (1973), 309-317.

ERMOLIEV Y., CHOR N.

1. Sur la minimisation de fonctions non dérivables. « Kibernétika », 1 (1967), 101, 102 (en russe).

FIKHTENGOLTZ G., KANTOROVITCH L.

1. Sur les opérations linéaires dans l'espace de fonctions bornées. Studia math. 5 (1935), 69-98.

FRÉCHET M.

1. Sur quelques points du calcul fonctionnel. Rend. Math. di Palermo 22 (1906), 1-74.

2. Sur les ensembles des fonctions et les opérations linéaires. C.R. Acad. Sci. (Paris) 144, (1907), 1414-1416.
3. Sur les fonctionnelles continues. Ann. Ec. Norm. 27, n° 3 (1910), 193-216.

FREUDENTHAL H.
1. Teilweise geordnete Moduln. Proc. Acad. Sei. Amsterdam, 39 (1936), 641-651.

FRIEDMAN V.
1. Sur la convergence des méthodes du type de plus forte pente. UMN 17, 3 (1962), 201-204 (en russe).

GATEAUX R.
1. Sur les fonctionnelles continues et les fonctionnelles analytiques. C.R. Acad. Sci. (Paris) 157 (1913), 325-327.

GAVOURINE M.
1. Über die Stiltjessche Integration abstrakten Funktionen. Fund. Math. 27 (1936), 254-268.
2. Méthodes analytiques d'étude des transformations fonctionnelles. Outch. zap. LGU. ser. mathem., 19 (1950), 59-154 (en russe).
3. Sur les estimations pour les valeurs et les vecteurs propres d'un opérateur perturbé. DAN SSSR 96, n° 6 (1954) 1093-1095 (en russe).
4. Equations fonctionnelles non linéaires et analogues continus des méthodes d'itération. Izvestia vouzov, Mathem., n° 5 (1958) 18-31 (en russe).

GONTCHAROV V.
1. Théorie d'interpolation et approximations des fonctions. Izd. 2-e, Gostekhizdat, M.-L., 1954 (en russe).

GRIBANOV Y.
1. Sur la mesurabilité d'une fonction. Izvestia vouzov, Mathem., n° 3 (1970), 22-26 (en russe).
2. Sur la mesurabilité des noyaux des opérateurs intégraux. Izvestia vouzov, Mathem., n° 7 (1972), 31-34 (en russe).

GROTHENDIECK A.
1. Sur les applications linéaires faiblement compactes d'espaces du type C ($K$). Canad. J. Math. 5 (1953), 129-173.
2. Produits tensoriels topologiques et espaces nucléaires. Mem. Amer. Math. Soc. 16 (1955).

GUELFAND I.
1. Abstrakte Functionen und lineare Operatoren. Math. zb. 4 (46) (1938), 235-286.

GUNTER N.
1. Théorie du potentiel. Gostekhizdat, M.-L., 1952 (en russe).

HARRIK I.
1. Sur un analogue de l'inégalité de Markov. DAN SSSR 106; n° 2 (1956), 203-206 (en russe).
2. Sur les approximations de fonctions s'annulant sur la frontière du domaine avec leurs dérivées partielles par des fonctions spéciales. Sib. mathem. journ. 4, n° 2 (1963), 408-425 (en russe).

HELLINGER E., TOEPLITZ O.
1. Grundlagen für eine Theorie der unendlichen Matrizen. Math. Ann. 69 (1910), 289-330.
2. Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten. Enciclopädie der matematische Wissenschaften, Bd II, C 13 (1927), 1335-1616.

HILDEBRANDT T. H.
1. Über vollstetige lineare Transformationen. Acta Math., 51 (1928), 311-318.

JAMES R.
1. A non-reflexive Banach spaces isometric with its second conjugate spaces. Proc. Nat. Acad. U.S.A. 37, nº 3 (1951), 174-177.
KADETZ M., MITIAGUINE B.
1. Sous-espaces complémentés dans les espaces de Banach. UMN 28, nº 6 (1973), 77-94 (en russe).
KAKUTANI S.
1. Some characterizations of Euclidean space. Jap. Jorn. Math. 16, nº 2 (1939), 93-98.
2. A generalization of Brouwer's fixed point theorem. Duke Math. J. 8, nº 3 (1941).
KALANDIA A.
1. Sur une méthode directe de résolution de l'équation de la théorie de l'aile et son application à la théorie de l'élasticité. Mathem. sb. 42 (1957) (en russe)
KANTOROVITCH L.
1. Sur une nouvelle méthode de résolution approchée d'équations aux dérivées partielles. DAN SSSR 2, nº 8-9 (1934), 532-536 (en russe).
2. Sur les espaces vectoriels semi-ordonnés et leurs applications à la théorie des opérations linéaires. DAN SSSR 4 (1935), 11-14 (en russe).
2a. Sur les propriétés des espaces semi-ordonnés linéaires. Comptes Rendus Ac. de Sc. (1936), 813-816.
3. Lineare halbgeordnete Raüme. Mat. sbornik 2 (44), 1937, 121-168.
4. The method of successive approximations for functional equations. Acta math. 71 (1939), 52-68.
4a. Linear operations in semiordered Spaces I. Mathem. sbornik 7 (49) (1940), 209-284.
5. Méthodes mathématiques d'organisation et de planification de la production. Izd. LGU, L., 1939 (en russe).
6. Sur le déplacement des masses. DAN SSSR 37, nº 7-8 (1942), 227-229 (en russe).
7. Sur une méthode efficace de résolution des problèmes d'extrémum pour une fonctionnelle quadratique. DAN SSSR 48, nº 7 (1945), 483-487 (en russe).
8. Sur un problème de Monge. UMN 3, nº 2 (1948), 225-226 (en russe).
9. Analyse fonctionnelle et mathématiques appliquées. UMN 3, nº 6 (1948), 89-185 (en russe).
10. Principe du majorant et méthode de Newton. DAN SSSR 76 (1951), 17-20 (en russe).
11. Sur les opérateurs intégraux. UMN 11, nº 2 (1956), 3-29 (en russe).
12. Sur certaines applications ultérieures de la méthode de Newton. Vestnik LGU, nº 7 (1957), 68-103 (en russe).
KANTOROVITCH L., AKILOV G., RUBINSTEIN G.
1. Etats extrémaux et commandes extrémales. Vestnik LGU, nº 7 (1967), 30-37 (en russe).
KANTOROVITCH L., GAVOURINE M.
1. Application des méthodes mathématiques dans les problèmes d'analyse des transports. Sb. stateï « Problemy provichenia effektivnosti raboti transporta ». Izd. AN SSSR, 1949, str. 110-138 (en russe).
KANTOROVITCH L., RUBINSTEIN G.
1. Sur un espace fonctionnel et quelques problèmes d'extrémum. DAN SSSR 115, nº 6 (1957), 1058-1061 (en russe).
2. Sur un espace de fonctions complètement additives. Vestnik LGU, nº 7 (1958), 52-59 (en russe).
KANTOROVITCH L., VOULIKH B.
1. Sur la représentation des opérations linéaires. Comp. Math. 5 (1937), 119-165.

KARPILOVSKAÏA E.
1. Sur la convergence d'une méthode d'interpolation pour les équations différentielles ordinaires. UMN 8, nº 3 (1953), 111-118 (en russe).
2. Sur la convergence de la méthode des sous-domaines pour les équations intégro-différentielles ordinaires. Sib. mathem. journ. 4, 3 (1963), 632-640 (en russe).
KELDYCH M.
1. Sur la méthode de Galerkine pour la résolution de problèmes aux limites. Izv. AN SSSR, ser. mathem., 6 (1942), 309-330 (en russe).
KELLEY J.
1. Banach spaces with the extension property. Trans. Amer Math. Soc. 72, nº 2 (1952), 323-326.
KIS O.
1. Sur la convergence de la méthode de collocation. Acta Math., Acad. Sci. Hung., 17, 3-4 (1966).
KOCHELEV A.
1. Méthode de Newton et solutions généralisées des équations non linéaires de type elliptique. DAN SSSR 91, nº 6, (1953), 1263-1266 (en russe).
2. Existence de la solution généralisée d'un problème de torsion élasto-plastique. DAN SSSR 99, nº 3 (1954), 357-360 (en russe).
KOLMOGOROV A.
1. Zur Normierbarkeit eines allgemeinen topologischen linearen Raumes. Studia Math. 5 (1934), 29-33.
KOROTKOV V.
1. Opérateurs intégraux dont les noyaux vérifient les conditions de Carleman et d'Ahiezer I. Sib. Math. journ. 12, nº 5 (1971), 1041-1055 (en russe).
2. Représentations intégrales d'opérateurs linéaires. Sib. math. journ. 15, nº 3 (1974), 529-545 (en russe).
KRASNOSSELSKI M.
1. Quelques problèmes d'analyse non linéaire. UMN 9, nº 3 (1954), 57-114 (en russe).
2. Sur quelques nouveaux principes du point fixe. DAN SSSR 208, nº 6 (1973), 1280-1281 (en russe).
3. Sur les équations opératorielles quasi linéaires. DAN SSSR 214, nº 4 (1974), 761-764 (en russe).
KRASNOSSELSKI M., KREIN S.
1. Processus itératif avec des résidus minimaux. Mathem sb. 31 (1952), 315-334 (en russe).
KRASNOSSELSKI M., TCHÊTCHIK V.
1. Sur un théorème de L. Kantorovitch. Tr. seminara po funkts. analizou, vyp. 3-4. Rostovsk. n/D. un-t, VGU (1960), 50-53 (en russe).
KREIN S., PÊTOUNINE Y.
1. Echelles des espaces de Banach. UMN 21, 2 (1966), 89-168 (en russe).
KREIN S., RUTMAN M.
1. Opérateurs linéaires laissant invariant un cône dans un espace de Banach. UMN 3, 1 (1948), 3-95 (en russe).
LERAY G. et SCHAUDER J.
1. Topologie et équations fonctionnelles. UMN 1, nº 3-4 (1946), 71-95.
LEVINE V.
1. Sur le problème de déplacement d'une masse. DAN SSSR 224, nº 5, (1975), 1016-1019 (en russe).
2. Problèmes d'extrémums avec des fonctionnelles convexes semi-continues inférieurement pour la convergence en mesure. DAN SSSR 224, nº 6 (1975), 1256-1259 (en russe).
3. Sur les théorèmes de dualité pour le problème de Monge-Kantorovitch, UMN 32, (1977) (en russe).
4. Problème de Monge-Kantorovitch sur le déplacement d'une masse. V kn.

« Metodi founktsionalnovo analiza v matematitcheskoï ekonomiké » M., « Naouka », 1977 (en russe).
LEVINE A., STRYGUINE V.
1. Sur la vitesse de convergence de la méthode de Newton-Kantorovitch. UMN 17, 3 (1962), 185-187 (en russe).
LIOUBITCH Y., MAÏSTROVSKI G.
1. Théorie générale des processus de relaxation pour les fonctionnelles convexes. UMN 25, nº 1 (1970), 57-112 (en russe).
LOMONOSSOV V.
1. Sur les sous-espaces invariants d'une famille d'opérateurs commutant avec des opérateurs complètement continus. Founktsionalni analiz i evo prilojénié, 7, nº 3 (1973), 55-56 (en russe).
LOZANOVSKI G.
1. Sur les espaces réticulés de Banach isomorphes. Sib. math. journ. 10, nº 1 (1969), 93-98 (en russe).
2. Sur certains espaces réticulés de Banach. Sib. math. journ. 10, nº 3 (1969), 584-599 (en russe).
3. Sur les fonctionnelles localisées dans les espaces de Riesz. Sb. « Theoria founktsii, founkts. an. i ich prilojénié », Kharkov, vyp. 19 (1974), 66-80 (en russe).
LOZINSKI S.
1. Les espaces $\tilde{C}_\omega$ et $\tilde{C}_\omega^*$ et la convergence des processus d'interpolation. DAN SSSR 59, nº 8 (1948), 1389-1392 (en russe).
2. Fonctions inverses, fonctions implicites et résolution d'équations. Vestnik LGU, nº 7 (1957), 131-142 (en russe).
LUXEMBURG W. A. J.
1. Notes on Banach Function Spaces. Proc. Acad. Sci. Amsterdam A68 (1965), 229-248, 415-446, 646-667.
LUXEMBURG W. A. J., ZAANEN A. C.
1. Notes on Banach Function Spaces. Proc. Acad. Sci. Amsterdam A66 (1973), 135-153, 239-263, 496-504, 655-681; A67 (1964), 104-119, 360-376, 493-543.
MYSSOVSKIKH I.
1. Sur la convergence de la méthode de Kantorovitch de résolution des équations fonctionnelles et son application. DAN SSSR 70, nº 4 (1950), 565-568 (en russe).
2. Sur le problème aux limites pour l'équation $\Delta u = k(x, y) u^2$. DAN SSSR 94, nº 6 (1954), 995-998 (en russe).
NACHBIN L.
1. A theorem of the Hahn-Banach type for linear transformations. Trans. Amer. Math. Soc. 68, nº 1 (1950), 28-46.
NATANSON I.
1. Quelques théorèmes non locaux des intégrales singulières. DAN SSSR 19, nº 5 (1938), 357-360 (en russe).
NEUMANN J. von
1. Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. Math. Ann. 102 (1929), 49-131.
NIKOLAÊV V.
1. Sur le problème de l'approximation des fonctions continues par des polynômes. DAN SSSR 61, 2 (1948), 201-204 (en russe).
NIKOLAÊVA G.
1. Sur la construction approchée d'une transformation conforme par la méthode des séries trigonométriques conjuguées. Tr. Mathem. in-ta im. V. Stéklova 53 (1959), 236-266 (en russe).
NIKOLSKI S.
1. Equations linéaires dans les espaces vectoriels normés. Izv. AN SSSR, Ser. mathem 7, nº 3 (1943), 147-163 (en russe).

PELTCHINSKI A., FIGUEL T.
1. Sur la méthode d'Enflo de construction d'espaces de Banach sans la propriété d'approximation. UMN 28, n° 6 (1973), 95-108.
PHILLIPS R. S.
1. On weakly compact subsets of a Banach space. Amer. J. Math. 65 (1943), 108-136.
POLIAK B.
1. Méthodes du gradient de résolution d'équations et d'inéquations. JVM i MF 4, 6 (1964), 995-1005 (en russe).
2. Une méthode générale de résolution des problèmes d'extrémum. DAN SSSR 174, n° 1 (1967), 33-36 (en russe).
POUGATCHEV B.
1. Sur la convergence des méthodes localement proches de la méthode de Newton-Kantorovitch. Tr. séminara po founkts. analizou, vyp. 7, Voronej, izd-vo VGU (1963), 130-136 (en russe).
RELLICH F.
1. Spektraltheorie in nichtseparablen Räumen. Math. Ann. 110 (1934), 342-356.
RIESZ F.
1. Sur les opérations fonctionnelles linéaires. C.R. Acad. Sci. (Paris) 149 (1909), 974-977.
2. Untersuchungen über Systeme integrierbarer Funktionen. Math. Ann. 69 (1910), 449-497.
3. Leçons sur les systèmes d'équations linéaires à une infinité d'inconnues, Paris, 1913.
4. Sur la décomposition des opérations fonctionnelles. Atti Congresso Bologna 3 (1928), 143-148.
5. Sur quelques notions fondamentales dans la théorie générale des opérations linéaires. Ann. of Math. 41 (1940), 174-206.
RUBINSTEIN G.
1. Quelques exemples de problèmes d'extrémums duaux. V kn. : Mathematitcheskoe programmirovanié. M., « Naouka », 1966 (en russe).
2. Dualité en programmation mathématique et quelques problèmes d'analyse convexe. UMN 25, 5 (155), (1970) (en russe).
3. Sur le déplacement d'une masse sur un compact. DAN SSSR 223, n° 3 (1975), 572-575 (en russe).
RVATCHEV M.
1. Sur le déplacement d'une masse. V. zb.: Optimisation, vyp. 9 (26), Novossibirsk, 1974, 203-208 (en russe).
SAMOKICH B.
1. Etude de la vitesse de convergence de la méthode de plus forte pente. UMN 12, n° 1 (1957), 238-240.
2. Méthode de plus forte pente dans le problème des valeurs propres d'opérateurs semi-bornés. Izv. vouzov, Mathem. n° 5 (1958), 105-114 (en russe).
SCHAUDER J.
1. Sur Theorie stetiger Abbildungen in Funktionalräumen. Math. Zetschr. 26, n° 1 (1927), 47-65.
2. Über lineare vollstetige Operatoren. Studia Math. 2 (1930), 183-196.
3. Über den Zusammenhang zwischen der Eindeutigkeit und Lösbarkeit partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Math. Ann. 106 (1932), 611-721.
4. Das Anfangswertproblem einer quasilinearen hyperbolischen Differentialgleichung zweiter Ordnung in beliebiger Anzahl von unabhangigen Veränderlichen. Fund. Math., 24 (1935), 213-246.
SÉDAEV A.
1. Sur un problème de G. Lozanovski. Troudy NIIM VGU, Voronej, vyp. 14 (1974), 63-67 (en russe).

SÉMÉNOV E.
1. Théorèmes d'immersion pour les espaces de fonctions mesurables de Banach. DAN SSSR 156, nº 6 (1964), 1292-1295 (en russe).
SIRVINT Y.
1. Ensembles convexes et fonctionnelles linéaires dans les espaces abstraits. Izv. AN SSSR, ser. mathem., 6 (1942), 143-170, 189-211 (en russe).
SMOLITSKI H.
1. Sur la sommabilité des potentiels. UMN 12, nº 4 (1957), 349-356 (en russe).
SOUDAKOV V.
1. Problèmes géométriques de la théorie des distributions probabilistiques infinidimensionnelles. Troudy MIAN 141, L., « Naouka », 1976 (en russe).
STEINHAUS H.
1. Additive und stetige Funktionaloperationen. Math. Zeitschr. 5 (1919), 186-221.
STÉPANOV V.
1. Cours d'équations différentielles. Uzd. 8-e, Fizmatguiz, M., 1959 (en russe).
TOULAÏKOV A.
1. Zur Kompaktheit im Raum $L_p$ für $p = 1$. Nachr. Ges. Wiss. Göttingen 1, nº 39 (1933), 167-170.
TROÏTSKAÏA E.
1. Sur les valeurs propres et les vecteurs propres d'opérateurs complètement continus. Izv. vouzov, Mathem., nº 3, (1961), 148-156 (en russe).
VERCHIC A.
1. Quelques remarques sur les problèmes de programmation linéaire en dimension infinie. UMN 25, 5 (1970), 117-124 (en russe).
2. Sur les travaux d'Orstein, les conditions de faible dépendance et les classes de mesures stationnaires. « Theoria veroïatnostei i iéo primenenia » 21, 3 (1976) (en russe).
VERTHEIM B.
1. Sur quelques méthodes de résolution approchée d'équations opératorielles non linéaires dans les espaces de Banach. UMN 12, 1 (1957), 166-169 (en russe).
VEXLER A.
1. Sur la localité des espaces de Riesz fonctionnels. Sib. math. journ. 12, nº 1 (1971), 54-64.
VLADIMIROV V.
1. Problèmes mathématiques de la théorie de transfert des particules. M., « Naouka », 1961 (en russe).
ZABREÏKO P.
1. Espaces idéaux de fonctions. Vestnik Iarosl. in-ta, vyp. 8 (1974), 12-52 (en russe).

# INDEX DES MATIÈRES

# INDEX DES SYMBOLES